

1 GAACCCAGTT GCTTCAGCGA GTCGAACTAC AGTTTTAACCC TCATCAAATA  
51 TGGCATCTCC CTTGCTTGC GCAGCAGGGA TGGAAAGAAAAT GTCACTTCT  
101 TTTTAAGCTA GCAAGCTTT TCTTTTCTT TTTCTTCTTC TATTAAAGAAA  
151 TTCTAATCAT GGATGCTTCT TCCGACCCCTT ATTGCCTTA TGACGGGGGA  
201 GGAGACAATA TTCCCCTGAG GGAATTACAT AAAAGAGGAA CTCATTATAC  
251 AATGACAAAT GGAGGCAGCA TTAACAGTTC TACACATTTA CTGGATCTT  
301 TGGATGAACC AATTCCAGGT GTTGGTACAT ATGATGATT CCATACTATT  
351 GATTGGGTGC GAGAAAAATG TAAAGACAGA GAAAGGCATA GACGGATCAA  
401 CAGCAAAAG AAAGAATCG CATGGGAAT GACAAAAAGT TTGTATGATG  
451 CGTGGTCAGG ATGGCTAGTA GTAACACTAA CAGGATTGGC ATCAGGGCA  
501 CTGGCCGGAT TAATAGACAT TGCTGCCAT TGATGACTG ACCTAAAGGA  
551 GGGCATTGCGT CTTAGTGCCTG TGTTGGTACAA CCACGAACAG TGCTGTTGG  
601 GATCTAATGA ACAACACATT GAAGAGAGGG ATAAATGTCC ACAGTGGAAA  
651 ACATGGGCAG AATTAATCAT AGGTCAAGCA GAGGGTCTG GTTCTTATAT  
701 CATGAACATAC ATAATGTACA TCTTCTGGC CTTGAGTTTG GCCTTCTTG  
751 CAGTTCCCT GGTAAAGGT TTTGCTCCAT ATGCCCTGTGG CTCTGGAATT  
801 CCAGAGATTA AAACATTTTT AAGTGGATTG ATCATCAGAG GTTACTTGGG  
851 AAAATGGACT TTAATGATTA AAACCATCAC ATTAGTCCTG GCTGTGGCAT  
901 CAGGTTGAG TTTAGGAAAA GAAGGTCCTC TGGTACATGT TGCCTGTTGC  
951 TGCAGAAATA TCTTTTCCTA CCTCTTCCA AAGTATAGCA CAACGAAGC  
1001 TAAAAAAAGG GAGGTGCTAT CAGCTGCCCT AGCTGCAGGG GTTCTGTTAG  
1051 CTTTGGTGC ACCAATTGGA GGAGTTCTT TTAGCCTGGA AGAGGTTAGC  
1101 TATTATTTTC CTCTCAAAAC TTTATGGAGA TCATTTTTTG CTGCTTTAGT  
1151 GGCTGCATTT GTTTTGAGGT CCATCAATCC ATTGGTAAC AGCCGCTG  
1201 TCCTTTTTTA TGTGGAGTAT CATAACCCAT GGTACCTTT TGAACTGTTT  
1251 CCTTTTATTTC TTCTAGGGGT ATTTGGAGGG CTTTGGGGAG CCTTTTTCAT  
1301 TAGGCAAAAT ATTGCCCTGGT GTCTCGACG CAAGTCCACG AAATTTGAA  
1351 AGTATCCCCT TCTGGAAGTC ATTATTGTTG CAGCCATTAC TGCTGTGATA  
1401 GCCTTCCCTA ATCCATACAC TAGGCTAAAC ACCAGTGAAC TGATCAAAGA  
1451 GCTTTTACA GACTGTGGTC CCCTGGAATC CTCTTCTCT TGTA  
1501 GAAATGACAT GAATGCCAGT AAAATTGTCG ATGACATTCC TGATCGTCCA  
1551 GCAGGCATTG GAGTATATTG AGCTATATGG CAGTTATGCC TGGCACTCAT  
1601 ATTTAAAATC ATAATGACAG TATTCACTT TGTCATCAAG GTTCCATCAG  
1651 GCTTGGTTCAT CCCCAGCAGTG GCCATTGGAG CGATCGCAGG AAGGATTGTG  
1701 GGGATTGCGG TGGAGCAGCT TGCTTACTAT CACACAGACT GTT  
1751 TAAGGAGTGG TGTGAGGTG GGGCTGATTG CATTACACCT GGCCTTTATG  
1801 CCATGGTTGG TGCTGCTGCA TGCTTAGGTG GTGTGACAAG AATGACTGTC  
1851 TCCCTGGTGG TTATTGTTTG TGAGCTTACT GGAGGCTTGG AATATATTGT  
1901 TCCCCTTATG GCTGCAGTC TGACCAGTAA ATGGGTTGGA GATGCCTTG  
1951 GCAGGGAAGG CATTATGAA GCACACATCC GATTAAATGG ATACCCCTTC  
2001 TTGGATGCAA AAGAAGAATT CACTCATACC ACCCTGGCTG CTGACGTTAT  
2051 GAGACCTCGA AGGAATGATC CTCCCTTAGC TGTCCTGACA CAGGACAATA  
2101 TGACAGTGGA TGATATAGAA AACATGATTA ATGAAACACAG CTACAATGGA  
2151 TTTCCGTCA TAATGTCAA AGAATCTCAG AGATTAGTGG GATTTGCCCT  
2201 CAGAAGAGAC CTGACAATTG CAATAGAAAG TGCCAGGAAA AAACAAGAAG  
2251 GTATCGTTGG CAGTTCTCGG GTGTGTTTG CACAGCACAC CCCATCTCTT  
2301 CCAGCAGAAA GTCTCGGCC ATTGAAGCTT CGAAGCATTC TTGACATGAG  
2351 CCCTTTACA GTGACAGACC ACACCCCAAT GGAGATTGTG GTGGATATT  
2401 TCCGAAAGCT GGGACTGAGG CAGTGCCTTG TAACTCACAA TGGGCGCTC  
2451 CTTGGCATTAA TAACAAAAAA AGATATCCTC CGGCATATGG CCCAGACGGC  
2501 AAACCAAGAC CCCGCTTCAA TAATGTCAA CTGAATCTCA CAGATGAGGA  
2551 GAGAGAAGAA ACGGAAGAGG AAGTTTATTG GTTGAATAGC ACAACTCTT  
2601 AACCTGAGGG AGTCATCTAC TTTTTTTTCC TCCTTACAA AAAAGAAAG  
2651 GAAATATAAA AGCCGGTTT TTGCAACATG GTTGGCAAAT AATGCTGGTG  
2701 GAATGGAGGA GTTGGTTGGG GAGGGAAAGG AGAGAGAAGG AAAGGAGTGA  
2751 GGTATTTCCC GTCTAACAGA AAGCAGCGTA TCAACTCCTA TTGTTCTGCA  
2801 CTGGATGCAT TCAGCTGAGG ATGTGCTGTA TAGTGCAGGC TTGCGCCTCA  
2851 ACAGAGATGA CAGCAGATC CTCGAGCACC TGGCCTGTTG CTCCAACATT  
2901 GCAAAGACAC ATTACATGTC CCTATTCTA GAGGGATTAC TTTGAATTGA  
2951 GCCATCTATA AACTGCAAG GTCTTGCCTT TTTTTTAAT CAAAAGTGT  
3001 CTGTTTAATT CATGAATTGT ATAGTTAACG ATTACCTTTC TACATTCCAG  
3051 AAGAGCCTTT ATTCTCTCTC CTCTCTCTC CTCTCTACTG  
3101 AGCTGTAACA AAGCCTTTTAAATCAGGTG ATCCTTTGAGCAGTCCTT

FIGURE 1, page 1 of 3

3151 TCTCATATTG AGATGTACTG TGATTTACT GAGGTTCAT CACAAGAAGG  
3201 GAGTGTTC TGTGCCATTA ACCATGTAGT TTGTACCATC ACTAAATGCT  
3251 TGGAACAGTA CACATGCACC ACAACAAAGG CTCATCAAAC AGGTAAAGTC  
3301 TCGAAGGAAG CGAGAACGAA ATCTCTATT GTGTGCCGTG TGGCTCAAAA  
3351 CCGAAAACAA TGAAGCTTGG TTTTAAAGGA TAAAGTTTTC TTTTTGTTT  
3401 TCCTCTCAGA CTTTATGGAT AATGTGACCG GGTCTTATGC AAATTTCTA  
3451 TTTCTAAAAC TACTACTATG ATATACAAGT GCTGTTGAGC ATAATTAAAT  
3501 AAAATGCTGC TGCTTGACA GTAAAGAGAA AAAAAAAA AAAAAAAA  
3551 AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA  
3601 AAAAAAAA AAAAAAAA AAAAAA (SEQ ID NO:1)

**FEATURES:**

5'UTR: 1-158  
Start Codon: 159  
Stop Codon: 2532  
3'UTR: 2535

www.ncbi.nlm.nih.gov

**HOMOLOGOUS PROTEINS:****Top BLAST Hits:**

		Score	E
CRA 18000005109762	/altid=gi 2599548 /def=gb AAB95161.1  (AF029...	1575	0.0
CRA 18000005109763	/altid=gi 2599550 /def=gb AAB95162.1  (AF029...	1573	0.0
CRA 18000005227216	/altid=gi 4762023 /def=gb AAD29440.1 AF14277...	1572	0.0
CRA 18000004989660	/altid=gi 4502869 /def=ref NP_001820.1  chlo...	1570	0.0
CRA 18000005231972	/altid=gi 8134363 /def=sp Q9R279 CLC3_CAVPO ...	1561	0.0
CRA 18000004989700	/altid=gi 6680948 /def=ref NP_031737.1  chlo...	1560	0.0
CRA 18000004978791	/altid=gi 1705905 /def=sp P51792 CLC3_RAT CH...	1560	0.0
CRA 1000685681515	/altid=gi 6634696 /def=emb CAA71072.2  (Y0994...	1559	0.0
CRA 18000004989661	/altid=gi 1705903 /def=sp P51790 CLC3_HUMAN ...	1558	0.0
CRA 18000005226296	/altid=gi 4753144 /def=gb AAB88634.2  (U8346...	1556	0.0

**EST:**

	Score	E
gi 10993825 /dataset=dbest /taxon=96...	1562	0.0
gi 10934924 /dataset=dbest /taxon=96...	1336	0.0
gi 10952244 /dataset=dbest /taxon=96...	1251	0.0
gi 12383593 /dataset=dbest /taxon=96...	1205	0.0
gi 6591096 /dataset=dbest /taxon=9606 ...	1170	0.0
gi 10251711 /dataset=dbest /taxon=96...	1104	0.0
gi 2321385 /dataset=dbest /taxon=9606 ...	1045	0.0
gi 5594360 /dataset=dbest /taxon=9606 ...	975	0.0
gi 5422132 /dataset=dbest /taxon=9606 ...	965	0.0
gi 10327969 /dataset=dbest /taxon=96...	963	0.0

**EXPRESSION INFORMATION FOR MODULATORY USE:****library source:**

gi|10993825 Neuronal precursor cells-teratocarcinoma  
gi|10934924 Whole embryo-mainly head  
gi|10952244 Neuronal precursor cells-teratocarcinoma  
gi|12383593 Small intestine-duodenal adenocarcinoma  
gi|6591096 Lung-small cell carcinoma  
gi|10251711 Breast-normal  
gi|2321385 Schwannoma tumor  
gi|5594360 Brain-tumor  
gi|5422132 Testis  
gi|10327969 Lung-large cell carcinoma

1 MDASSDPYLP YDGGGDNIPL RELHKGTHY TMTNGGSINS STHLLDLLDE  
51 PIPGVGTYDD FHTIDWVREK CKDRERHRRI NSKKKESAWE MTKSLYDAWS  
101 GWLVVTLTGL ASGALAGLID IAADWMTDLK EGICLSALWY NHEQCCWGSN  
151 ETTFEERDKC PQWKTWAELI IGQAEGPGSY IMNYIMYIFW ALSFAFLAVS  
201 LVKVFAPYAC GSGIPEIKTI LSGFIIRGYL GKWTLMIKTI TLVLAVASGL  
251 SLGKEGPLVH VACCCGNIFS YLFPKYSTNE AKKREVLSAA SAAGVSVAFG  
301 APIGGVLFSL EEVSYYFPLK TLWRSFFAAL VAAFVLRSIN PFGNSRLVLF  
351 YVEYHTPWYL FELFPFILLG VFGGLGWAFF IRANIAWCRR RKSTKFGKYP  
401 VLEVIIVAAI TAVIAFPNPY TRLNTSELIK ELFDCGPLE SSSLCDYRND  
451 MNASKIVDDI PDRPAGIGVY SAIWQLCLAL IFKIIIMTVFT FGIKVPSGLF  
501 IPSMAIGAIA GRIVGIAVEQ LAYYHHDWFI FKEWCEVGAD CITPGLYAMV  
551 GAAAACLGGVT RMTVSLVVIV FELTGGLEYI VPLMAAVMTS KWVGDAFGRE  
601 GIYEAHIRLN GYPFLDAKEE FTHTTLAADV MPRRRNDPPL AVLTDQDNMTV  
651 DDIENMINET SYNGFPVIMS KESQRILVGFA LRRDLTIAIE SARKKQEGIV  
701 GSSRVCFQAQH TPSLPAESPR PLKLRSILD SPFTVTDHTP MEIVVDIFRK  
751 LGLRQCLVTH NGRLLGIIITK KDILRHMAQT ANQDPASIMF N (SEQ ID NO:2)

**FEATURES:**

**Functional domains and key regions:**

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION

N-glycosylation site

Number of matches: 5

1	90-93	NETT
2	364-367	NTSE
3	392-395	NASK
4	587-590	NMTV
5	598-601	NETS

[2] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE

CAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

1	24-27	KKES
2	330-333	RRKS
3	331-334	RKST

[3] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE

Protein kinase C phosphorylation site

Number of matches: 8

1	22-24	SKK
2	333-335	STK
3	529-531	TSK
4	613-615	SQR
5	631-633	SAR
6	642-644	SSR
7	658-660	SPR
8	709-711	TKK

[4] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 13

1	27-30	SAWE
2	34-37	SLYD
3	92-95	TTFE
4	93-96	TFEE
5	105-108	TWAE
6	217-220	STNE
7	249-252	SLEE
8	383-386	SLCD
9	589-592	TVDD
10	666-669	SILD
11	674-677	TVTD
12	679-682	TPME
13	709-712	TKKD

[5] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 18

1	49-54	GLASGA
2	53-58	GALAGL
3	72-77	GICLSA
4	88-93	GSNETT
5	189-194	GLSLGK
6	206-211	GNIFSY
7	234-239	GVSVAF
8	240-245	GAPIGG
9	245-250	GVLFSL
10	310-315	GVFGGL
11	313-318	GGLWGA
12	314-319	GLWGAF
13	408-413	GVYSAI
14	447-452	GAIAGR
15	491-496	GAAACL
16	541-546	GIYEAH
17	638-643	GIVGSS
18	692-697	GLRQCL

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	99	119	1.810	Certain
2	182	202	2.131	Certain
3	233	253	1.398	Certain
4	256	276	1.019	Certain
5	290	310	1.770	Certain
6	321	341	0.797	Putative
7	361	381	2.093	Certain
8	400	420	1.539	Certain
9	473	493	1.739	Certain
10	496	516	1.218	Certain
11	540	560	1.568	Certain
12	570	590	0.975	Putative

**BLAST Alignment to Top Hit:**

>CRA|18000005109762 /altid=gi|2599548 /def=gb|AAB95161.1| (AF029346)  
chloride channel protein 3 [Homo sapiens] /org=Homo  
sapiens /taxon=9606 /dataset=nraa /length=818  
Length = 818

Score = 1572 bits (4026), Expect = 0.0  
Identities = 764/765 (99%), Positives = 764/765 (99%)

Query: 27 GTHYTMNGGSINSSTHLLDDEPIPGVGYDDFHTIDWVREKCKDRERHRRINSKKKE 86  
GTHYTMNGGSINSSTHLLDDEPIPGVGYDDFHTIDWVREKCKDRERHRRINSKKKE  
Sbjct: 54 GTHYTMNGGSINSSTHLLDDEPIPGVGYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87 SAWEMTKSLYDAWSGLVVLTGLASGALAGLIDIAADWMTDLKEGICLSSLWYNHEQCC 146  
SAWEMTKSLYDAWSGLVVLTGLASGALAGLIDIAADWMTDLKEGICLSSLWYNHEQCC  
Sbjct: 114 SAWEMTKSLYDAWSGLVVLTGLASGALAGLIDIAADWMTDLKEGICLSSLWYNHEQCC 173

Query: 147 WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206  
WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA  
Sbjct: 174 WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTTITLVAVASGLSLGKEGPLVHVACCCG 266  
PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTTITLVAVASGLSLGKEGPLVHVACCCG  
Sbjct: 234 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTTITLVAVASGLSLGKEGPLVHVACCCG 293

Query: 267 NIFSYLFPKYSTNEAKKREVLSAASAAGVSAGVAPIGGVLFSLEEVSYFPLKTLWRSF 326  
NIFSYLFPKYSTNEAKKREVLSAASAAGVSAGVAPIGGVLFSLEEVSYFPLKTLWRSF  
Sbjct: 294 NIFSYLFPKYSTNEAKKREVLSAASAAGVSAGVAPIGGVLFSLEEVSYFPLKTLWRSF 353

Query: 327 FAALVAAFVLRISINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 386  
FAALVAAFVLRISINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA  
Sbjct: 354 FAALVAAFVLRISINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 413

Query: 387 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSLCD 446  
WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSLCD  
Sbjct: 414 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSLCD 473

Query: 447 YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 506  
YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI  
Sbjct: 474 YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 533

Query: 507 GAIAGRIVGIAVEQLAYYHHDFWIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566  
GAIAGRIVGIAVEQLAYYHHDFWIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL  
Sbjct: 534 GAIAGRIVGIAVEQLAYYHHDFWIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593

Query: 567 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 626  
VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL  
Sbjct: 594 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 653

Query: 627 AADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSNGFPVIMSKESQLVGFALRRDLT 686  
AADVMRP RNDPPLAVLTQDNMTVDDIENMINETSNGFPVIMSKESQLVGFALRRDLT  
Sbjct: 654 AADVMRPRRNNDPPLAVLTQDNMTVDDIENMINETSNGFPVIMSKESQLVGFALRRDLT 713

Query: 687 IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSIILMSPTVTDHTPMEIVVD 746  
IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSIILMSPTVTDHTPMEIVVD  
Sbjct: 714 IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLSIILMSPTVTDHTPMEIVVD 773

Query: 747 IFRKLGLRQCLVTHNGLGIITKKDILRHMAQTANQDPASIMFN 791  
IFRKLGRLRQCLVTHNGLGIITKKDILRHMAQTANQDPASIMFN  
Sbjct: 774 IFRKLGLRQCLVTHNGLGIITKKDILRHMAQTANQDPASIMFN 818 (SEQ ID NO:4)

>CRA|18000004989660 /altid=gi|4502869 /def=ref|NP\_001820.1| chloride  
 channel 3; ClC-3 [Homo sapiens] /org=Homo sapiens  
 /taxon=9606 /dataset=nraa /length=820  
 Length = 820

Score = 1567 bits (4013), Expect = 0.0  
 Identities = 764/767 (99%), Positives = 764/767 (99%), Gaps = 2/767 (0%)

Query: 27 GTHYTMNGGSINSSTHLLDEPIPGVGYDDFHTIDWVREKCKDRERHRRINSKKKE 86  
 GTHYTMNGGSINSSTHLLDEPIPGVGYDDFHTIDWVREKCKDRERHRRINSKKKE  
 Sbjct: 54 GTHYTMNGGSINSSTHLLDEPIPGVGYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87 SAWEMTKSLYDAWSGLVVLTGLASGALAGLIDIAADWMTDLKEGICLSSLWYNHEQCC 146  
 SAWEMTKSLYDAWSGLVVLTGLASGALAGLIDIAADWMTDLKEGICLSSLWYNHEQCC  
 Sbjct: 114 SAWEMTKSLYDAWSGLVVLTGLASGALAGLIDIAADWMTDLKEGICLSSLWYNHEQCC 173

Query: 147 WGSNETTFFEERDKCPQWKTWAELIIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206  
 WGSNETTFFEERDKCPQWKTWAELIIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA  
 Sbjct: 174 WGSNETTFFEERDKCPQWKTWAELIIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTTITLVLAVASGLSLGKEGPLVHVACCCG 266  
 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTTITLVLAVASGLSLGKEGPLVHVACCCG  
 Sbjct: 234 PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTTITLVLAVASGLSLGKEGPLVHVACCCG 293

Query: 267 NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSVYYFPLKTLWRSF 326  
 NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSVYYFPLKTLWRSF  
 Sbjct: 294 NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSVYYFPLKTLWRSF 353

Query: 327 FAALVAAFVLRISINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 386  
 FAALVAAFVLRISINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA  
 Sbjct: 354 FAALVAAFVLRISINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 413

Query: 387 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSLCD 446  
 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSLCD  
 Sbjct: 414 WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSLCD 473

Query: 447 YRNDMNAKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 506  
 YRNDMNAKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI  
 Sbjct: 474 YRNDMNAKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 533

Query: 507 GAIAGRIVGIAVEQLAYYHHDFIFKEWCEVGADCITPGLYAMVGAACLGGVTRMTVSL 566  
 GAIAGRIVGIAVEQLAYYHHDFIFKEWCEVGADCITPGLYAMVGAACLGGVTRMTVSL  
 Sbjct: 534 GAIAGRIVGIAVEQLAYYHHDFIFKEWCEVGADCITPGLYAMVGAACLGGVTRMTVSL 593

Query: 567 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE--EFTHT 624  
 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE EFTHT  
 Sbjct: 594 VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEEFEFTHT 653

Query: 625 TLAADVMPRLNDPPLAVLTQDNMTVDDIENMINETSNGFPVIMSKESSQLVGFALRRD 684  
 TLAADVMPR RNDPPLAVLTQDNMTVDDIENMINETSNGFPVIMSKESSQLVGFALRRD  
 Sbjct: 654 TLAADVMPRRNDPPLAVLTQDNMTVDDIENMINETSNGFPVIMSKESSQLVGFALRRD 713

Query: 685 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSIIDMSPFTVTDHTPMEIV 744  
 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSIIDMSPFTVTDHTPMEIV  
 Sbjct: 714 LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSIIDMSPFTVTDHTPMEIV 773

Query: 745 VDIFRKLGRLQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791  
 VDIFRKLGRLQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN  
 Sbjct: 774 VDIFRKLGRLQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 820 (SEQ ID NO:5)

>CRA|1000685681515 /altid=gi|6634696 /def=emb|CAA71072.2| (Y09941)  
     putative chloride channel ClC-3 [Xenopus laevis]  
     /org=Xenopus laevis /taxon=8355 /dataset=nraa  
     /length=791  
     Length = 791  
  
 Score = 1559 bits (3993), Expect = 0.0  
 Identities = 745/791 (94%), Positives = 771/791 (97%)  
  
 Query: 1 MDASSDPYL PYDGGGDNIPLR ELHKRGTHYT MNGGSINSSTHLLDLLDEPIPGVGTYDD 60  
         MD SSDPYL PYDGGGDNIPLR+LHKRGTHYT+TNGG+INS+THLLDLLDEPIPGVGTYDD  
 Sbjct: 1 MDISSLDPYL PYDGGGDNIPLR DLHKRGTHYT VTNGGAINSTHLLDLLDEPIPGVGTYDD 60  
 (SEQ ID NO:6)

**Hmmer search results (Pfam) :**

Model	Description	Score	E-value	N
CE00039	CE00039 chloride_channel	1671.9	0	1
CE00420	E00420 CLC	1288.1	0	2
PF00654	Voltage gated chloride channels	1172.4	0	1
PF00571	CBS domain	78.1	7e-20	2

**Parsed for domains:**

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00654	1/1	71	622 ..	1	621 []	1172.4	0
PF00571	1/2	645	690 ..	11	54 .]	31.4	5.8e-07
CE00420	1/2	32	697 ..	1	729 [.	1174.4	0
PF00571	2/2	726	778 ..	1	54 []	47.4	2.2e-11
CE00420	2/2	722	791 .]	867	942 ..	110.6	6.5e-32
CE00039	1/1	60	791 .]	1	804 []	1671.9	0

1 AATTCTATAAC AAATATAATT ATATAGATAT ATATTACATA TACACACAAT  
 51 TGTTTATCTT TAAAAATAAT TCAAATATGG CTACAAAATC TTTACAATAT  
 101 GAAGCATTGT CAGTATTTAT TTTACCGGGA CGATTTCCCC CATCAGTGAG  
 151 TGCTGACTGT CATTTCATT CTTTATGATC AAGTTGTAGA TCAGGAAAAA  
 201 CAAGTTAAGA GAGTGCCTAC AAATACCGGG AAAACTTGTG GATAGATTT  
 251 CATTTTTAT GTAAAGACAT ATAAGAACAT GAATGGTATA AAAACAAAAT  
 301 CCTTTATAAA TGCCATACAA TTATATATT AGAAAAAATTA TATGGTGGTA  
 351 AAACATATAA AAGAACACCA CACTCCAAA TTTACATTGA GCTAATTAG  
 401 TACAGTTAGC CTTGTCAA GCTTCCCTTG TTTAAAAAAA CTATTGGCTC  
 451 AGTGTGCAGG AAGGAGCATA GGAGAAAAAA TTGCCAAGAA TATTGAAAAA  
 501 ATACAGAAAA TAAAGAAAAA AATCACCTAC TATCCTATCA AAAATTAA  
 551 TAGCTAGAAT CAGGATAAGA TAGAATATTCT GTGTGGCAGT AATTCTAGTC  
 601 TATATTCTT CCCTGGAACC CTGTCCTCCA AATTCAGGT GAGATTTAT  
 651 AAGAAGCTCT GTTTATCTGA GATTTAAAT ATAAGAACCT GATTAAACCT  
 701 ATACAGTTTT TAAAAAGAC CCTAAATAAG TAAAATTAG TACTCCACAA  
 751 ATTGAAGAGA ATTTCTCTCT TCTCTTTACT GCCCTCTGAG TTTTCTCTT  
 801 CCTTCTCTCA CCTCCAATTTCATGAAAC ACTTTCAGTT CGAGTGGACC  
 851 TTAGAGATTG TCTCATTCAA TACTTAGGA AAACAAATT TATAGAACCC  
 901 TTGAGTTCTG TGGAATTGCT TCTAATGAAC AACACCTTT GTTGGTTTG  
 951 TTGTTTAGTG ACACGTGTA ACAGGCATT CAGGAGGAGA ATCTCCCAGT  
 1001 CTAGAGGAAT CCTCTCAGAG GTAGCTATAA AATATTGAAC TCTGATCTC  
 1051 AATAAGCATT GTGCGGTTTT TGTTTTGTT TTAATGACA GTTTAAACA  
 1101 AGAAAGTTGC TTTATTTCTG AACTTCATAA AAATTTCTAT TAAAGAGACA  
 1151 ATTTCTGAAT TTTATAACAA TTTCTAGAAC AGTTGAGTAC CTCACTTTGA  
 1201 GACACATTTC TGCTAAAAGT TAAAACACAA AAACCCCTAT GAGATAAAAT  
 1251 AGGAAGCTAG TAGAGATAGG AAAGTCCTCT GCTTAGTAAA CCTCTTTTT  
 1301 GCGTAGTTA GACACATACA ATAGTAAAGT TACTTAGTAC GTTGATAGTT  
 1351 TTCTTCTCC TCAAAGCTA CAATGCTTTA CTAGCTAGTT CCTTCAAGAA  
 1401 AGGAAACAAG AAGCCGCTGG AGGAGATTGG TGAGTGGGAT AAAACACTAT  
 1451 TCAACTCTTC AGTTATTCGG TTTTTAAATC CTCAATGAAA GGCTGCTGTA  
 1501 TTATAGAGTA TTTTTTTTTT ATTTTTAATA GACTTAGAAC CAAGTTCTT  
 1551 GAGAAACCTT TGGCATATTG TAGTTTTTT ATGGCTATGA CTCACATGAC  
 1601 ATTACTGTAT AAAACTAGTA CATTCTCTCG TAAAACACAA CAAACTTACT  
 1651 AGAGTGCTGC TCTCATTTCATGAAAC ACTTTAGA AATGAAAAAG GGCATTGCT  
 1701 GCATCAAAA TTTCTTTTTT ACATCTCTGT ATTACTTTT CCCCTTTATA  
 1751 TTTATCTAA ACCAAAAGA AATAATGTTT CTATTGTTT ACTGTAGTTA  
 1801 CCACTGATGC TACCGAAGCT GTATTGTGAG TGTTCAAAA TTCTCAAACC  
 1851 AGTTTGTGT GTTGTACTTG GAGCTTAGTC ATTGTCAACAC GTAGCAGGAC  
 1901 CTGATTAAGA AGGCTGTGCG GCCTCTAACG CTTGCTAGAT GTAGCCACT  
 1951 AGCAACCAGG CTGCAATAAT TTCCCTTGA TGACATCATC CACTGTGGAA  
 2001 GAACCCAGTT GCTTCAGCGA GTCGAACATC AGTTTAACC TCATCAAATA  
 2051 TGGCATCTCC CTTGCTTGC GCAAGGGGA TGGAAAGAAAT GTCACTTTCT  
 2101 TTTTAAGCTA GCAAGCTTTT TCTTTTCTT TTTCTTCTTC TATTTAAAAA  
 2151 TTCTAATCAT GGATGCTTCT TCCGACCCCTT ATTTGCCTTA TGACGGGGGA  
 2201 GGAGACAATA TTCCCCCTGAG GGAATTACAT AAAAGAGGTA ATACTATCCC  
 2251 CTTGCTGTGA ATTCTCTGTT GGTATGTTT GCATGCGGCT GGGCGGTCCCT  
 2301 CTAGCTAAA CTGGTTCTCG TTTGCTCTT AAATACTGCA GTACGTTGTT  
 2351 TAGTTGCCCT GGGTTGTTAG TAAGGGAAA ATGCAACCTT CTGAATGGTT  
 2401 GTGTAGCCAT CCCTGATTGT TTTCTCTGTG CAGATTAGTA CTGCTTCAGA  
 2451 TCACGTCGGG CTCCGACTCC ATCTTCTGCA TGAAAATCTT CTTTCTAACT  
 2501 CTGAAAATGA ATTAATCTGC TTTTACAGCC AACTAAAGTC GTGTTGGTTG  
 2551 GCATCTAAA AGTAATGTTT TTCTCCTTC AGAAAACCTTA CATTTCCTT  
 2601 AATTTCACACA GAGAAATCAG GTGCCTATGT ACCATTATAT TTTAGCTGCT  
 2651 GCCAATTACC ATGTAGATT TACACCACAA AGTAAATTTA TAGCAAAGC  
 2701 TTTACCTACA TTTTAGAACA TTTTAAATG ATAGTAAAGA TGAATAATT  
 2751 CTATATTAAT ACTTTTTATT TAATATGTT TTCGGCTGAG TAACATACTA  
 2801 CATTGTCTTC CACAGGTATC TTGTGAAATT TGATATGATA AAACACATT  
 2851 GACTAAATGT CAGAAAAAAAT AATATTGGTT TGTAAGAAC AGAAGAGCAC  
 2901 CCAGCATGCC TGTAATCTT TTGGCAGGCA CTTCCCTCAGT CTCCTTAAA  
 2951 TTAATTGCTGAT GTTAATTACT ACCCTTTTT TCATTTTGT TTAATTGCTT  
 3001 ATTCAAGAAAAA CAGACTGGTC GACATTGTT GTCCCTAGAAA AAAATTGAAC  
 3051 TTCAAGAAAAA ATCTCTTAGC TTATGTGACT TCATTTTGTG GCCACATTAG  
 3101 TTTGAATTAC TGCAATGATAT TATAAACTCA CCTTATGATT TAACCCAAAC

FIGURE 3, page 1 of 27

bioRxiv preprint doi: <https://doi.org/10.1101/2023.09.11.553021>; this version posted September 11, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

3151 TTTTATTTGT AAGTATATAA GGAAGTAATA ATGTTTTCT AATATAATTA  
3201 GCCTGCTTTA TTTAAAATAT ACTTTGTGTT CTGATAACAC TTTTTTTTA  
3251 GTATTAAGTT CCACTATAAT TTAAACATTA TAATGTATTG AACAAATGTC  
3301 TGTTGGTTGC ATTGTGTCAG CTACACACTA TTGAGGGTC TGAACAGTTG  
3351 TAGCATTATT TATCTTGCAAG TATTCTGTAG TTAGTAAAAA CTTGCTTTT  
3401 ACATTTGAG AAAAGCTGTG TAAGGATCAT GTTACATACA TTGCTTTTC  
3451 TCTTACAGAG TTACCTTCTT AATAAAATTT TGATATATGT GTATATGTAT  
3501 ATGTTAGAAC ATTTGGAAGA AATATCTAAA AGCATAAAGA AGAAAATAAT  
3551 TTCTTGTAAAT CACACCACCC AGAGCTTTT AAATTTTTT TCTTAATGTT  
3601 ACGATCATAA ATTCTTCTAT TTCTTGTGTT CTGATTATCA GTTTTCTGGT  
3651 AAGGAGTTCT TAAACAGGAA AGCAAGGTGA ATGAATAGTG ACTGTTCAAA  
3701 TGTACATTA TTTGCTAATC AGTAATTAAA CTGTAACAAAGA AGACAGACTG  
3751 TATTTTCTC ATGCTATTAC AACATTGGT TTGTAATGAT GATAGATCAG  
3801 AATACCTGGG CTTCAGAAAT TTAAATTCCCT TTGTAAGC TTAACAGTCT  
3851 TTGACAGAAC TTACTTATGG ACTGCTTAG TGAAAATAT GCAAATAATA  
3901 AGAAATAAGT CAAAACCTTAT GTGAGAGTAG GCATGGTTAC TGATATTACC  
3951 TAAACGTAAG CTTTTTATTCTT CTATTATACT TTCAATAATA ATCCTTAAG  
4001 AATCTTGCTT AGGATCTAAA TCAGTCCCAC TCTTGGCAGC TCAAATAGGT  
4051 TCTTTATCCC TTGATGAGAC TTATTCTATT AATATAAGTC ATTGTTATTT  
4101 GAAAGTAACA TTGTTGATGT GTAGTAGAGA TAAGTCAGTT ATTAGGCTTT  
4151 CGTGAAGTGT CTGTATTAC TCAAACATAC TGTAGTATCC TAGTGTCTAT  
4201 GCGTAAGATG TTATTTTTG TCCATAATTG ATGACCTGTT GTAGCCATGG  
4251 GTCAACACAA TGGAAATTGAT GGAGACAGGC AGCTAACAAA TCGAAAAAAC  
4301 TGAATCAGCT TCCCTGTGAG GAAGAACAAA ACTATAATGA TTAAAATGAA  
4351 TCTTCAGCCT GATAGTGAAG AGGCAGATAA AGTATAAAAT TGTGAAGGAT  
4401 ATCAATAAAAG TAAACATGGA TCTGTTAGT AAATCCCTGA GTGCTATAGC  
4451 CAAGGATTAC CTTTGTGAG TAAATTGAAT TTAATACTAC TTTTCAAGGC  
4501 GAGATGGTAA ATGGTGAAGC TTCTTATTAA AGTAAATAAT GTCAAGTCTG  
4551 GAAGTATAAG TAGATTCAAA TTAGAATTAG TTGATATAC TATTGATAGA  
4601 TTAGAAATTAA AGATGACATT TCAGAAATAG CCATCTTTAG GGGTAGATTT  
4651 CCTATATAGA ACAATCAAG CTCTCTCAAAT GTGCTCTTC CTTTTTTATC  
4701 AGGAAAAAAAG ACTTGGCTTA TCTGGACTGT TAGTTTACA CTTTTTCTTC  
4751 TTAAATTGTT CAAGATGTTT AAGTAGTTT AGAGGTCAA TTTCTTCTT  
4801 CTACCAACCC TTATTAATGG ATTGATTCT TTGGGCCCTG AGCCTCCATT  
4851 TACTCCATGA GGGGCCCTTA ACAATTATTT AAATNNNNNN NNNNNNNNNNN  
4901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
4951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
5001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
5051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
5101 NNNNNNNNNNN NNNNNAAAAT AGTAATATTA ATAATAGTTA ATATTTATTA  
5151 GAATTCTCTG TTAGCTGGAT ACTGTCCTAAGTGGTTTT TTTGTTGTTG  
5201 TTGTTGTTGT TGTTGTTTC TTAAGAGAGA GGTATCACTT TTTCACCCAG  
5251 GCTGGAGTGC AGTGGAGTGA TTATAGCAAA TGAGCCTTG AACTACTGGG  
5301 CTTAGATCCT CCGTCTCACCT CTCCTGGTA CCTGGGACTG CAGGCTTGCA  
5351 ACACCTTGCC TGGCTAATTAA AAAAAACAAA ATTTTTTTTT TTTTAGGGA  
5401 GAGTCTCACT ATGTTGTCCTA GGCTGGCTC CAACTCCTGG GCTCAAGCAA  
5451 CCCTCCTGCC TTGGCCTCCC AAGTAGCTGA GATTACAGGT GCGAGCCACT  
5501 GTGCCCTGGCT TGTTCTAAGT GCTTTATGTG TATGAAATTAA TTAAATCCT  
5551 CATCACAAGT TTATGAAGTA GGTACTGTTA TAATCCCCAT TTTCTAGTTG  
5601 ACAAGACTGA GGTAAGGAAT TTGTAAGGAA AAGTCAGAAT TCCATCCAGA  
5651 TATTGGCTC ATACTTTAAT CATGAGGCTA AACTGCTTCT CTCTACACGT  
5701 ATCTTCATAG TAACTTGTGT TTTAAGTCTG GTAGAAGCAT AAGAAGTTA  
5751 AACACAGACA GAATCCTGTG GAAGTTAGTA AATTCTAGT GAACGATAGA  
5801 AATGATAGAA ATCTCTTCTT CCCCCAAAGT CCCAAGAACAA GATTAGTCTG  
5851 CTTTGACAA GTGTTATCAA AGTAGACTGT TCTCACATAC ACGGGGGACT  
5901 CAATAGGGCA TTCTGGTGG ATATAATAAAT ATGAGTAAAT GCGATAACAG  
5951 GAGGAAATGC CTAGTGTGTT GCTCTGGAT TAGTTTTGAT ACAACAAAGG  
6001 CAGCTTGTGTT GTGAGTCAGT AGAGAGGGTA GTGTAGAAAG GTGGAAGTTG  
6051 GAAGAGTGGC AGATCCTAGA GGACTAATGA TGGGCTTAAAC CCACAAAAAG  
6101 TGTCGCTTTG CCATTGAAAT AAAAGTTGG GGTCTTATTT TTTCAATT  
6151 CTCCCTGAAA TTATTTCTT ACATTCTTAC GCTCAGCAGT GTATCTAAAT  
6201 AAAGCTTTT TGGGTTCTA TTATAATAGA GTGTTGTTCC TTTTTCTTCC  
6251 CTTGAAAAG TATCATTCTT TGCACATTAT TTGAAAATCC AGGTGTTATA

6301 TGATATTCTT ATTGCCAGAG GGACATTCTG CAGGCTCTT GTAAAATGAT  
6351 TTTAGGATTC AGATACTTAT TATATTTTA TTGGCCCTAA TATTTTATCC  
6401 AACTAGAAAA TTAAACCTCT TCTTAAAAAT TAATCCATCT AAGTGTCTGT  
6451 AAATTAAAGG ACAACTAAA GATTCTTAT TTGGTGTCAAG AAACCTCTTG  
6501 TTTCTACAAAC AGTAGTATAA AACAAAGCCT GTTTTAAAT GTACTTTCC  
6551 CACAGTATCT GAATTCTAAA TCTTCATAA AATCTGGITC ATATTACTAC  
6601 CTCTAGCTTG ATTTCTAAA AATAGCTGAC ACTTTAGTAT GGTTAATTT  
6651 ATGCCATCTC ATGGCTGTG AGAAATGCTT TGTATCAAGA TTTCCGAGTG  
6701 TGAACAGATT TCCTGCCGCA TTGATTAAGT TTGTAATTTT GGCTATTTC  
6751 CCAGCATCGA GTTTCTGCT TTGCGTTTAT GCAGGAGACT GGTAGTTAA  
6801 ATTGAACCTT AAGGTTTGT TTCTGTTTT TAAGTTAACAA TATGTTAAT  
6851 TTCTAGTTTC TTTGTAGCCC TTTGCAACTT TAATTAGGTC ATAAAATGGA  
6901 TTTACTCTAG TTTCTCTAAC AAATTTTATA AATTATGAA ATATGAAATT  
6951 TAGCAAATTT TATAAACCTT TTTATTCTG TATTGTACAG CTCATCATAT  
7001 TTGAGACAT ATAATTGAA TGTGGAACCTT GTTCCAATT ACACAGATGT  
7051 CTTAATATCC ACCTTATCAT CTCTAACTAA AGGATGTGGC TTTTTATTT  
7101 TGAGGTGGCA ACAGAACAGA AAAGAAAACA GTGAATTGAG TAATGGGCTT  
7151 AGTATTGCTG CTGCCTGGTT GTGTATCTT GGTAACCTTC TTTGAGATT  
7201 GGCATTAACCT TGCAAGTCTT TGCAAGTTAG ACAGTTAAAT ATGACTGAAT  
7251 GGCTGAACAA ATTTTAATAG CGTATGCTTC TTTTTTGCTA TTTTATTAC  
7301 CAGTAGACAT TTAATTGACC ACCTGCTAA TGTGAGGCAC TATTCTGCC  
7351 ATTACCTTTT TAATCTTGTG TTTGGAGTCT GCTAACATTG TGGAACCTCC  
7401 ACTATCAACT TAGAACGTTT ACTTCCCCT CCCTTACCAAG GATGGCCATT  
7451 TCTTATCAGT AGGGTCACAG AGAGAGAAAA AAAAACCAT CTGGGGCTAG  
7501 ACTTCCTGCT CTTAACATAC AGAACGAAAT AGGTTGTGAA GGAATACATA  
7551 GTATTGGAA TTTCTGCCTC TTCCTCCAT AATTTTTTA AAAAGGTCA  
7601 TATTTTTAT GTGTGTCTTA TGTAAACAGTA ATCTGCATTA TGAACCTAAA  
7651 TGACGAGGAT CACCAATTCA CATCTTGGA GATTGATCAC AGAGGTAATA  
7701 AGTAACCTT TTTAAATAAC TATATGATC ATTTCATG TAAAACATT  
7751 ATTTGGATAA ACCCCTTTGA GAAAAGGCTT AGGCTCCTGC CAGTGTCACT  
7801 GTGATTTTA CTAATAAGCT CAGTTAACCG CGCAGCAATT AAGGTTGTGT  
7851 TGTGTTTTTT TTTTAAGTT CAGTTCAGCA AATATATGTG GAAAGCTGT  
7901 GGGTAAAATT ATATTTGTAT TTTTGGGAA GCACACAATT TTATTAATGC  
7951 CTATTTTTT CTAGTTCACT GTTTGTCAAA CTTCAAGTTT TAACATGTTG  
8001 ATCATGAAAC CAGTTGACTT GTGACCAAGTA TTTAAAAGG AAAGATTTAA  
8051 AAAACAAAAT AAAATATCA TATATACCA GTAGTAAGAG TAAGCATTTG  
8101 TTACTAAACT TTGGTTTTAT TTAAGTACAT ATCTATATAC TATGTCAGTG  
8151 AGAACATTT CTCCACTTCA TGTTGAAAA ACATTTCAAA AGCTAAGAAA  
8201 AAGTTGAAA ACCTGTTGTG AAGTACACCT GGGGTAAGG TACACCTGT  
8251 GGCATAAGAT GTCGGGAAACA ACTGAGGGTA AGAATGGGA TGCATTACTA  
8301 TCGTAAACTT CTGCTAAAGC ATAAGGATGT GAGTGCTGGG AGCAAAGCAG  
8351 TGCTCACCAC TTCTGCAATT TTCTATTGCA GCATTTAAA TAATATGGG  
8401 AAAAGTGGAC TGCAACCAAA GGCAAGAGG GATGGTGATG GTGAAGGGTA  
8451 AGATTGTATT TATTGTCCAA AGGCTAAGTG CATAACATA TGTGTTGGG  
8501 AGAAGGCATC ACGTAATAGT TCTTAACCTA CTCTGAGAGA AGGTTGTCCA  
8551 CATTCTTAA AGTATACATG TAAACCAACA ATGAAATTAT TTTAGTGA  
8601 TGAGAATCAA AGTGTAGAG TTTGAATCCC TGTCTACTA CTTGCTAGCG  
8651 GTGTGACCTT GGGCCTGTT AACTCTGAC ACCTTGTTTT CCAAATTAT  
8701 AAAGTGGAGA TAATAATATC TGTCACTTG TGTGTTGTG AGGATTATAT  
8751 GAACTAATAT ATGTAATGTC CTGAGAACAA TGTCTGGTAC ACATTAAGTT  
8801 AATTAAAATT AGCTGTTCTT ACTGTTATTA TTAGACATGA GCTAGATAAC  
8851 AGTGGCCTCT ACATGTGAAA GATTATTTA ATTCTGATGT AGTCAGTTT  
8901 ATCTATTTTT TTTATTTTG TCCCTTTGC ATTGATGTCA TATCTAAAAA  
8951 ACCTGCCTAA CTCAGGATCA CAAAATTTA CTCCGTATT TTATAATTT  
9001 AGCTCTTCTT ATCTAGGATC CATTCTTGC TAATTTTTAT ATATGGTGTG  
9051 AGGTAGGGT ACGGTTTCAT TCTTTGCAC GTGAATAGCC AGTTGTCCA  
9101 GCATCATTTA TTCAAAAGAC TATTCTTCC TCACTAGAAA AAATATTCT  
9151 TTAAAGAATA ATGAATCCTT TTTTTTTCT TTTAACCGC TGTTACTCAG  
9201 TTGGAAAAAG ATAATGAAT AATTAAAGT AATTTCCTA CAGGTAATT  
9251 TAAGTCCTTA TGTTTAGATT ACACATATTA GGAAATAATG GATTTGTATT  
9301 CCATAGGTAT GCTTGATCTT TATAAAGTTC CCTGTCTCTG GAAAACCTAA  
9351 AATAAGGCAA ACAATCTTC TTAGTAGAGT TATTTTACA AGAAAGTTGC  
9401 AAGCCAGTTT TAGTCATCG ATTGGATAAT TTTCTGTGCT TGCTGGAGGT

9451 ATTCAGTAT TGGTAATACC TGAACATG A GGATGCATGA ATGATGCATT  
9501 TTAGGAATTT GTTCTGTGT CCATACCGAG CATAATGAAT TAAGTTATCT  
9551 GTTAAAAATA CAGGATTTTT GCTCAATATA CAGTTGTAGA AGAACTCATT  
9601 GTCCAAATT TTAAGACTTT TTTTTCTTTT TTTTTTGAG ATGGATCTCG  
9651 CTCTGTCGCC CAGGGTGGAG TGCACTGCC CAACCTCCAC TCAC TGCAAC  
9701 CTCCACCTCC AGGGTTCAAG TGATTCTGCT GCCTCAGCTT CCCGAGTAGC  
9751 TGGGGACTAC AGGCATATGC CACTATGCC GCCTGATTTT TTTTAGTAGA  
9801 GATGGGGTTT CACCATATTG GCCAGGCTGC TCTGAACTC CTGACCTCGT  
9851 GATCCACCCG CCTCAGCCTC CCAAAGTTCT GAGATTACAG GTGTGAGCCA  
9901 CCGCGCCCCG CGAGACATT TTTTTTTTT TTTTTTTTT GCTGTCTTG  
9951 TCATATTGTT AGTCTTTGG TTAAGCGATA TTATAACTTA GTCAATATGAG  
10001 TAATATAATG CAACATGCTG AATTGTGTGT GTGAGAGGGG GTTGTCTTG  
10051 GTTTGTTATT TGTTTTTAA ATAGAGATGA GATCTCACTG TGTTTCCAG  
10101 GCTCCCTTGA ACTCCTGGGC TCAGATGATA TAGCCTCCTG CCACAGCGTC  
10151 CTGATTAGCT GGGACTACAG GTGTGCACCA CTACACGTGG CTTTCTGAT  
10201 GAAATTTAA ATACCCAAAT ATTTGAGCAG AAATAATAGC TTGTGTTTAT  
10251 TGTTTTCTA CTATCTGTCA AGTATAGTAT TAAATGTTT ACATAATTG  
10301 TCTCCAGTCC ACATACAATA CTCTAGTAGA AGTGGGTAAC AAAACCAAGG  
10351 TACTCAAAGA GTTAAATAAG TAACTTGCAC TGGATCACAG AACTAACGGG  
10401 AGGCAGGGCT GGAATTGAC TCTAGGTCTT TCTGACCTCA AAGTGCAGTA  
10451 AAGTCATGGA ATTTCTCTAC TAGGCCACCT GGAAGAAAAG TGATCTTTT  
10501 TCCAGTCTT TTTGTTACTG TTTTCAGCC AGGAGATAGT AGAGTTAGGT  
10551 AGTAAATAG TAGTCACTGG CATCCGGTAG TCAGCCCTCC AAAAAAGTTT  
10601 TTGATTTTTT TTTTTTTTT TGTCTAAAC TTGGAAGCTA CTAACCTTCA  
10651 GGTCAACTT TCTTATCATC CAAGAGCTGG ATATTTAGGT AGCAGAAACT  
10701 ATGGAATTAT CCTAAGTCCT CTTGAAGCTT CAGCTGTTAA ATTAATTGG  
10751 TTCTGATTAA CACTGTGCTC AAGATTACA TTTCTAGGAG CCACAGTTG  
10801 ATTGGTCTAA CTTGGATCTA TGTGTTTCT TTAGCTGGGG AGGAGAAGGT  
10851 ATCTTGATTG ATACCTTCAC CAGGACTGCA TGCACTGAGG GACAGAAGTT  
10901 TCCTTAAAT AATTGGGTTC TGTTATAGGA AGAAGGGAA GGAGATACCA  
10951 AGTGGGCAAA ACAATCAGGT TCTATTACAT AAATAATAAA CCTAATGTGA  
11001 CGATAATAAA TGGATAATAT GATTATTTA AGTTTGGAAA TATACTGGT  
11051 TATTAGTATT GGATATCTGG TAGTGGGGTT GGAGAAAAAG TCGAGAATAA  
11101 GAAAAGACTT AAAATCGTA AAATTAACCTG GAAAAGAGGA TGGCTGAGCA  
11151 GATACATATA TGTAGATGAA TGTTCATATA GGCAAAACCAA CCTGAAGATT  
11201 TGTTAAATT GTAGTATGTA GCCAGGTGTG GTGGTGTCTTG CCTGCAGTCC  
11251 CAACTACTTG GGAGGCTGAG GCAGGATGTG TGCTTGAGCC TAGGTTGAG  
11301 GCTACAGTGA GCTATGTTTC CACCACTGCT TTCCAGCCTA GGTGGCAGAG  
11351 CAAGACCCCA TCTCTAAAAA AATAAAGTAA ATGAATAAA TTATAATATG  
11401 TTATGACAAA TATAGTTATC TGAAGTCACA GAAAATGTGC ATGTGCATT  
11451 AATGATGTGA ATAATTTTT AGGAAGATG AATAAAAAAA TCAACCTTTA  
11501 AGTGTGGCTA GTATGATCTT ACCTGTATCT CACTTATAGA AAATATAAAA  
11551 GGCTGAAGCC AGTCACCAAGT TTAATAGTTC TAAACCTCTTG TTTACTTGAT  
11601 TCCCTTTTTT CTCCCTCCCCA GCAATCCTCA TATAGTTAGG TAAAGTTGGT  
11651 TCTTCATCAG GCTTGTGCA GAAACCCCTA AGCCTTTTTA CTTAAAGCTT  
11701 TTTGAAACCC AGAAACCCAT CTTTGAAATT CAAAAGTTT GACTGTTATT  
11751 AGTCTTTTG TATGTTGTT GGCGCATAA ATGTCCTCCTT TTTATGAACA  
11801 GAGAAGTGTG TGTTAATATA CTTTGCAC TTTTGATGG GGTTGTTGT  
11851 TTTTTCTTG TACATTGTT TAAGTCTT GTAGATTCTG GATATTAGAC  
11901 CTATGTCAGA TGGATAGATT GCAAAAGTTT TCTCCCATTC TGAGGTTGC  
11951 TTGTTCATTC TGATGATAGT TTCTTTACT GTGCAGAAGC TCTTTAGTTT  
12001 AATTAGATCC TATTGTCCTG TTTTGGCTTT TGTCGCCATT GCTTTGGTG  
12051 TTTCACTCAT GAAGTCTTG CCAGTGCCTA TGCTGTAAT GGTATTGCCT  
12101 AGGTTTCAT GTTTGGGT TTTACATTTA AGCCTCAAAT CGATCTTGAG  
12151 TTAATTTTG TATAAGGTGT AAGGAAGGGG TCCAGTTCCA GTTTCTGCA  
12201 TATGGATAGC CAGTTTCCC AGCACCAATT ATTAAATATTA AATAGGGAAT  
12251 CCTTTCCCCA TTACTTGTGTT TTGTCAGTT TGCTGAAGAT CAGATGATTG  
12301 TAGATGTGTG GTGTTATTC TGAGGCTTT GTTCTGTTCC GTTGGTCTGT  
12351 ATATGTGTTT TGGTACCAAGT ACTATGCTGT TTTGGTTACT GAGCCTTGTA  
12401 GTATGTTTG AAGTCAGGTA GTATGATGCC TCCAGCTTIG TTATTTTGC  
12451 TTAGGATTGT CTTGGCCATA CGGGCTCTTT TTTGGTTCCA TATGAAATT  
12501 AAAGTAGGTT TTTCTAATT TGTGAGGAAA GTCAATGGTA GCTTGATGGG  
12551 AATAGCGTTG AATCTATAAA TTACTTCGGG CAGTATGGCC ATTTTCATGA

12601 TATTGATTCT TCCTATCCAT GAGCATGGAA TGTTTTCCA TTTGTTGTG  
 12651 TCGTTTCTTA TTTCCTGGG CAGTGGTTG TAGTTCTCCT TGAACAGGTC  
 12701 CTTCACGTCT CTTTTAAGTT GTACTCATCA TCACTGATCA TTAGAGAAAT  
 12751 GAAAATCAAA ACCACAATGA GATGTCATCT CATGCCAGTC AAATGGTGT  
 12801 TATTATAAA AGTCAAAAAA GAATAGATGT GGTAAGGCT GTGGAGAAAT  
 12851 AGGAATGCTT TTACACTGTT GGTGGAGTG TAAATTAGTT CAACCATTGT  
 12901 GGAAGACAGT ATGGCGATTCT CAAGGATC TAGAACCAGA AATACCATT  
 12951 GACCCAGCAG TCCCATTACT GGGTGTATAC CCAAAGGATT ATAAATCATT  
 13001 CTGCTATAAA GACACATGCA CACGTATGTT TATTATAGCA CTATTACAA  
 13051 TAGCAAAGAC TTGAAACCAA CCCAAAAAGC CATCAATGGT AGACTGGATA  
 13101 AAGAAAATGT GGCACATATA TACCATGGAA TACTATNNNN NNNNNNNNNN  
 13151 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13201 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13251 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13301 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13351 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13401 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13451 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13501 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13551 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13601 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13651 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13701 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13751 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13851 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 13951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14101 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14151 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14201 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14251 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14301 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14351 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14401 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14451 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 14501 NNNNTAAAAG ATACATCCTT TATTGATGCG TAAGATGAAA TCGAGAGGTG  
 14551 AAATTGGATA TACTGTTGCT TTTAAAAAT TTTAACATAT ATGTAATTTT  
 14601 TTGTTACTTAT CTCATTTAG CCTATATAAG TTATATATAT TTTGTTGTT  
 14651 TGGTTGTTG TTTGTTGAT GATGGAGTCT TGCTCTGTCA CCCAGGCTAG  
 14701 AGTCAGTGG TGCAATCTCG GCTCACTGCA ACCTCGCCT CCTGCATTC  
 14751 AGCGATTCTC CTGCCTCAGC CTCCTGAATA GCTGGGATTA CAGGCACCTG  
 14801 CCACCGCGCC CAGCTAATT TTTTATTT TAGTAGAGAC AGGGTTTCAC  
 14851 CATCTGGCC AGGCTGGTCT TGAACCTCTG ACCTGTGAT CCATGTGCC  
 14901 TAGCCTCCC AAGTGTGGG ATTACAAGCG GGAGCCACCG CGCCCGGCTG  
 14951 TAAGTTATAT CTTACACAAA TCTAGGTTTC ATTCAAGAGAA TTATATGCAA  
 15001 AGAACACAGT CAATAGGATT ATTTTAAAGC TATTGTTATT GTTAGAAAAC  
 15051 ATAATACCTT TAAAATTCCT TTTCACATTA GAAATATAGT GGCTTCTCCC  
 15101 CAGTTTAGGA TAGAAATT TCTTTCTTC TCCTTCTTTA TACTATTCTG  
 15151 ATTTGCATGT TTGACAGAAC AAATTATAAG AGAAAATATT TGAAATGTCA  
 15201 CATACTAAAG TAAATGTTG AATGTTGAA AATTTCTGG TTTTCAGAGA  
 15251 TTTTGAATTG CTGAATCGTT GTGAAATTAA AGATGTTGAG TAGTTCCAC  
 15301 AGAGTAATTA TTTGAAAGTC ACTGAAAGCA AGACACATGC CTAATGTAAA  
 15351 TGTTTATTGC ACTACTGTAC CTTTTCTAC CTCATAAAA TGAGAATAGC  
 15401 AGTCGTACT TTTCCACTTC GTCATTGTA AGTCTTTGCA GAAATTCTATA  
 15451 TTTGTTTGC TTATTATCTT CACGCTGTAAG ATAGCTTGAA AATTCTTAA  
 15501 GTGGGGCTAG CGATGTATTA TGGATACATG TTAAGTGGTA TAGAAATTTC  
 15551 ACTTTTTTTT TTTTGCATAA AGAGTAACAA GACCAAGTAGT CCATATTCT  
 15601 TCAGCTCTAC CCAGAGAAGG GCAATGTAGG AGGGAAAATG AAGTTGCAA  
 15651 AATATTTCAT AGTAGGCTTT TTCTTAAAGT AACCTCAGAC TTACAGAAGT  
 15701 TTAAGAATAG TACAAAGAAT CCCCATATAC CTGTCACCCCC AATTCTGAA

15751 ATATTAATAT TTTACCACAT TTGTTCATTA TGTCTGTATT CTCCAAGTAC  
15801 GATATATGCC ATTATATGTA ATATGTAGCA TTTTATATAG ACATAGGCA  
15851 TGTATGCACT ATATATTTTT TTCTGAGCCA CATAAAGAGT AAAACGCAGA  
15901 CATGACGTGC TTTTACTCCT AAATACTTCA GTGTGTGTAT TCCCTCAAGA  
15951 AAGGGCATTT TCTTCTGTAT AGCTACCGTA CACTTCTACA CTTTTCAAAA  
16001 TCAGAACATT TACATTGATA CCATACTATG ACATGATCTG CAGACCATT  
16051 TCCAATATGC CAGTTGTCCT ACTGTGTCTT TTAGTACAAA AGAAAAAAAGT  
16101 TTTTTTCCTT GGTCTAGGAG CTAATCCTGG AGCACATGTT ACATCCTGTT  
16151 GTTTTAATCT AGAACCGTTC CTCAGTTCTT TATCTTTCAT AACCTTGACA  
16201 TTTTTGGAGA GTACAATCCA TATATTGTC AGAATTTCCC TTAGTTGGG  
16251 TGTGTCTGGT TTTCCCTTAT AAGATTCTT TTATGCATT CTGGCCAGAG  
16301 TACCACAGAA GTACTGTATA TCTTACAGA AAGCCTAAGT GGCATTTGCA  
16351 TTTTCTAAAT GATCAATTAAAT ATTATATAT GGAAAGCAGA GTCAGAGATT  
16401 CTCACATATG TCAAGATATT ATAAGTATTCT CGTTTATATT TATTCTCAA  
16451 TTGCTTTTC TCAAGAAAAT TTGTGGCCTT TCAGCTAGCT TTTCAAAGTG  
16501 GAAGTTACTA CATAACATTA GGATGGGAGG GGTGGGGAAG AGCTTTATTA  
16551 AAGCTTTAAG ATTGAGCTTT TGAGTATGTG TTGTATGTAA ATGAAAGTGG  
16601 GCATTGATGC AGGGATTGGG CCTTTAAACC TTTGGCCAAG AATGGTATCA  
16651 ATTATTATTA TTATTATTT TTGGAGTACT TCTGCTAAA CACTGAAATC  
16701 AGTGTGCCAC TCTCCTTTA GAAGTTTAC ACCTTTCCAA GGTACACTTT  
16751 TTTTTTGGA GACGAGTTT GCTCTGTCGC CCAGGCTGGA GTGCATTGGC  
16801 GCAATCACAG CCCACTTCAG CCTCTGTTTC CCAGACTCCA GCAGTCCTTC  
16851 CACTTCAGCC TCCCGAGTAG CTGGGATTAC AGGTGCACAC CACCATGCC  
16901 AGCTAGTTT TGTAGAGATG GGGTTTGCC CATGTTGCC AGGCTGGTCT  
16951 CCAACTCCTG CGCTCAATCT ATCCGTCTC CTCAGCCTGC CAAAGTACTG  
17001 GGATTACAGG CGTGGGCCAC CACTCCGGC TTCCAAGGCA GGCATTTAAA  
17051 TGTAATAAAAT AGGGAGATAA GCAAGAACCC TGTGGACCT GGTAGAAGCA  
17101 AACATTATTA AGTACTATTA CGTTGTTAA AATATTAGCG CCTTCTATAT  
17151 TCATGTCCTC CCAGAATTAT CAAAAAAACCT ACTCTATAGT TTATTTGGCT  
17201 TATATCTCAG GAGTAATAAA ATTAGTTAAT AGTATTGGCA TCGTGGTCT  
17251 TTGTTGATTCT CTCCTTATC CCACCCCAAG TTGATTTCAC ATGATCTCTT  
17301 GATCTAGTCT AAGAATGTTT ATAGTGATTA CGAGAAGTTC AGATTCTGGC  
17351 TTTAACATAT ATAATTTGTT TTTAACATCTG AAACCAAAGA GAATGAGTTT  
17401 GTTTAAACTA GAAAGATGGC AAGAGTAGTC TGGGAATTTT GTTCCATTCC  
17451 TAAAAAGTCC TATAAAAAAA TAAACATATC TTGTTGTTTA TTTTTACAAAT  
17501 TTTTTAAAC ATTAGTACAG AGTGCCACTT CTTATATTCT ATATCAAATA  
17551 ATGAGCTACA TTTTCAATAA TAACCTCTGA GTAATTGTTG GCATTAAAAT  
17601 GCTGCATTAC AAAATAATTG GAGGATATAA TTTATAATCA CTTATGCTAA  
17651 AATCACCTAT TTGAAATTAT GTATGAGGTT TTCAAAGTTT ATAGTGCCTT  
17701 GGAAAAAATT TAAATGTTTC TTTGTTTATG TATCTTTATT ATAAGCTGTA  
17751 GCATATATCA TGTAGTTGTC AAGGATGCTG ATAGATACTT AATATTTAAA  
17801 GGAGACTTGT CAAAGTTAG CTGTCCAGGA CTAGAACCTG GGCCTTTGG  
17851 TAACAGCTCA TTGCTCTATT TACTTAAATG ATGATTGGAT TCGTTAGAAT  
17901 TTCTCTATTT TCATAGCTGT CTCTATGGTT CTATGAAAAT ACTGTGTGTG  
17951 TGCTTATACA TATATGTATA CCTGTAAGTA CAAAGTAGAA AATGAAAGTT  
18001 CATTCTCTGC TTTTGACAAT TGTAATCCCC AGAGATAACC GTTATTAAATA  
18051 TGTTGTCTCA TGTTGGTCA TACTGTTTTC TCTGTATTCT GTGTATTACT  
18101 GTATAAAATT TACACAGTAA TTTGCATATT AAAATGCTG GTCTACACCT  
18151 GGCCCTTTT TAAAAAACTGC AATTATTAT GGCAATTNTT TTATACCAAGT  
18201 ATATATTGAT CAACCTTATT CTTTTAACT GCTGCATTTC ATTCAATTAC  
18251 AATAGATGAG ACATTTCCAT TGGTTGAAT TTTTCAGTAT TACAGATAAT  
18301 GGTCAAATTA AATATTTAAG CTTTGTGCA CTTGAGAAT TAATTCCTAG  
18351 ACATAGAACCTT CTTATATTGT GATAGGTATT TCCAAATTTC TTCCCAAAT  
18401 GTTTGTATCT CTTTACTTCC ACTCTCAGGT CTAATAATTTC TCACTTGAT  
18451 TATCATATTT CTTACCCAGC CTGTTTTTA CACTCTAAC TCTTTTCTT  
18501 TTCTTTTTTT TTTTGAGACA GCATCTGCT CTTGGCCCGG TTGAAATGCA  
18551 GTGGCACGAC GACCAACCTG GGCTCAAGCA ATTCTCTCAA CTTAGCCTAC  
18601 TGAGTAGCTG GGACTACAGA CACATATCAG CAGCCAGC ATTTTTTTT  
18651 TTTTTTTTT GGATTTTTAG TAGAGATGAG GTTTGCCAT GTTGCAGAAG  
18701 CTGGTCTCAA ATTCTGTAGC TCAAGCAATC CACCCATCTC AGCCTCCAA  
18751 AATGCTGGGA TTACAAGCGT GAGCCACTGC ACCTGGCCCA AAAGCTCTT  
18801 TTCTAAATAGC AATATAAAATT GTCTTTACAG GACTATACTC ATATATGTTT  
18851 CTTCTTCAG AAATAGGTGT TAAGTGTATC TAACATGGAA TGTAGTAC

18901 TAATTCTCAT TGTGAAACCA TAGCCTAATT TATTCATAT TACAATTAA  
 18951 AATTCTATT TTTAGGAAG TTTCTTAGA TTAATCCGCC TAGTCCAGG  
 19001 TGCTACAGTC CCAAGATTC TTTCTTTTA ACAATTAAA TATAGGTAAAC  
 19051 ATGACTAGAA TTGTAAGCAA AGAATATTGG AACCTTGAA CTTCAGTATT  
 19101 TGAACCTTAT TTTGAAATAT AATTGTTAT ATTAAAAAA TATTATAATA  
 19151 TATTGCACCT GGAAGTTAGG GGCAGTTTT TTTAATTCTC TTTGTATCTG  
 19201 CTACACTGTA AAGTGCTATT TATGTAAGAA ATTCTTAATA GAAGTCTTCA  
 19251 GTTGTAAAGT CTGCTGTACA GACTTAGAT CAGGGATTGG CAAACTATGA  
 19301 GCCATGTGCC AAATCCTGCC CTTCACCTGT TTTGAAATA AAGTTTATC  
 19351 AGAACACATT CAGACTCATT CATGAACATA TTGCTATGA TTTATTTCT  
 19401 GCTACTATGG CAGAATTGAG TTGTTGCAAC TGTGTTGCAT CAAAGCCTA  
 19451 AAATATTTAC TCTCCTGGCT CTTTCCAAC CGGTTTAAAGA TTATGAGCAC  
 19501 TTTGCATTA TTATGTTTT GTTTCTTTC TATAGCACAC AGTAAGATGT  
 19551 TCTGCCACA TTGTCATAA TTTATGGTT TATTCAAGGA TTTATGCAAG  
 19601 TGTAGCTGCA AGAAAAAAAC CTAGAAGTGA ACTTGCTAGG TTGAAGAGCA  
 19651 TCTGTGTATG TTAAATTTTG TTAGCTTCG CCTTCCCAA GGGATTATTC  
 19701 CATTTCATAC TTAAACTACT AATTGTTGA TAGGACTTCT TTCTCCATAG  
 19751 CTTTGCTAAA TTAATGCATT CACACACTTC ATCTTTACTA ATCTGATAGA  
 19801 GGGAAATGAT ATTGTTGATT TGATTGCA TTCTTTTAT GTGTTAGCTT  
 19851 GAGCTTATTT TCATATTTAA AAGCCAATTG TATTCTTTT TCTTGAGCTA  
 19901 TCTTTAATG TCCTTCCTGA TACATTCTG AAGTCTGTGA TACTCATATA  
 19951 AGATATATGG TGAACATGTG TCAAAGATT ATTGACTCT AATGAGGGAA  
 20001 CCCGCCTGAT GACAAGGCTG ATTGAGAAGA GGATGTGTGA GATGAAGTGT  
 20051 ATATCATCAG TGAAAGAAAG CAAATTCTT CAGGGCAAAA ACAAAACCAC  
 20101 AACTCTAAGG GTTATTGTTT CTACTGGACA GAATTCAATT GCATTTTACC  
 20151 AGATAAAAAT TACTATTTTCA AATTATCTT TTACAAATCA TTTTCTAATT  
 20201 TTACAGAGTC TATTCCCTAA TCAGTAGTAA ATAGTCTTCA AAATTCTCCG  
 20251 CAGCGTCAGG TGACTATTAT GCAGGCTAAT TGTGACACT CGGGCTTGAC  
 20301 TTTAAGAGAA CATGCCATAA TCTTTGGCC TTACTTCCAA GTTTGGATA  
 20351 ATTTTCTTA ACACATTTT CTCTAATTGC AATGATTTCAGTGTATTAA  
 20401 TTCTTTTTT TTAAATTTT TTACTATTTA TTGATCACTC TTGGGTGTTT  
 20451 CTCGGAGAGG GGGATTTGGC AGGGTCATAG GACAATAGTG GAGGGAAAGGT  
 20501 CAGCAGATAA ACATGTGAAC AAAGGCTCT GGTTTCTTA GGCAGAGGAC  
 20551 CCTGCGGCCT TCCACAGTGT TTGTTGCTCC GGGTACTTGA GATTAGGGAG  
 20601 TGGTGTGAC TCTTAATGAG CATGCTGCCT TCAAGCATCT GTTTAACAAA  
 20651 GCACATCTTG CACCGCCCTT AATCCCCTTA ACCCTGAGTT GACATAGCAC  
 20701 ATGTTTCAGA GAGCAGGGGG TTGGGGGTTA GGTATGGAT TAACAGCATC  
 20751 CCAAGGCAGA AGAATTTTC TTAGTACAGA ACAAAATGGA GTCTCCTGTG  
 20801 TCTACTTCTT TCTACACAGA CACAGTAACA ATCTGATCTC TCTTTTCCCC  
 20851 ATATTTCCCC TTTTCTATTG GACAAAATG CCATCCTCAC CATGGCCCGT  
 20901 TCTCAATGAG CTGTTGGGTA CACCTCCCAG ACAGGGTGGC GGCCAGGCAG  
 20951 AGGGGCTCCT CACTTCCCAC ACTGGGGCGGC CGGGCGGAGG CGCCCCCCAC  
 21001 CTCCCAGACG GGGCGGCTGC CGGGCGGGGG CGCCCCCCCAC CTCCCAGACT  
 21051 GGGTGGCCGG GCGGAGACGC TCCTCACTTC CCAGATGGGG CGGCTGCCGG  
 21101 GCGGAGGGGC TCCTCACTTC TCAGATGGGG TCGCGGCTGG GCAGAGGTG  
 21151 TCCTCACCTC CCAGACAGGG TGGCGGCTGG GCAGAGACGC TCCTCACCTC  
 21201 CCAGACGGGG CAGCCGGGCA GAGGCCTCC TCACATCCCA GAGGGGGCGG  
 21251 CCGGGCAGAG GCGCTCCCCA CGTCCCAGAC GATGGGCAGG CGGGCAGAGA  
 21301 CGCTCCTCAC TTCCTAGACG GGATGGCGGC GGGGAAGAGG CGCTCCTCAC  
 21351 TTCTCTAGATG GGATGGCGGC CGGGAAAGAGG TGCTCCTCAC TTCTCTAGACT  
 21401 GGGGGCCGG GCAGAGGGGC TTCTCACATC CCAGACGATG GGCAGTCAGG  
 21451 CAGAGACGCT CCTCACTTC TAGTACAGGG TGGCGGCGG GCAGAGGTG  
 21501 CAATCTCAGC ACTTCGGGAG GCCAAGGCAG GTGGCTGGGA GGTGGGGTT  
 21551 GTAGCGAGCC GAGATCACGC CACTGCACTC CAGCTGGGC AACATTGAGC  
 21601 ACTGAGTGAG CGAGACTCCG TCTGCAATCC CGGCACCTCG GGAGGGCGAG  
 21651 GCGGGCAGAT CACTCGAGGT CAGGAGCTGG AGACCCAGCCC GGCCAACATG  
 21701 GCGAAACCCC GTCTCCACCA AAAAACACAA AAACCACTCA GCGCTGGCGG  
 21751 CGCGTGCCTG CAATCCCAGG CACTCGGCAG GCTGAGGGCAG GAGAATCAGG  
 21801 CAGGAAGGGTT GCAGTGAGCC GAGATCGCGG CAGTACAGTC CAGCCTCGGC  
 21851 AACAGAGGGGA GACCGTGGAAG AGTGGGGAGAC GGAGACGAGG GAGAGGGGG  
 21901 GACCGTGGAAG AGCAGGGAGGT GGAGACGAGG GAGAGGGAGA GGGATTATTT  
 21951 CTGTATGACT TAATAATGAA TTTCTAAGAG GTCACTTAGC TCACTGTTGT  
 22001 CTCTTCTAAA ACATACTCAT CTTTCTTTT CTCTTCTGTAA GGAACCTCATT

22051 ATACAATGAC AAATGGAGGC AGCATTAAACA GTTCTACACA TTTACTGGAT  
 22101 CTTTTGGATG ACCAATTCC AGGTGTGGT ACATATGATG ATTTCCATAC  
 22151 TATTGATTGG GTGCGAGAAA AATGTAAGA CAGAGAAAGG CATAGACGGG  
 22201 TAAGTGTGTT TAGTAAAAAT TTTTAAAAAC ATAGTCATA ATTAGATCTT  
 22251 TTAATAATAT ATTTCTGCCA ATGATCTCAG GCTGCCAAAT GTTACATTT  
 22301 AATATAAGTA AATGTCATA TTTCATATGT GGTACATGTT TTTTCTTTT  
 22351 TCTATGTTA ATTTTTAG TTTACTTATA CCCTGTAACT TTCCAGAAAG  
 22401 GATTTCAGGT AGCTAAAAAA CAAAGAAATA CAATAAGAAG ACAAAATAAG  
 22451 AAGGAAAGGG AAAAATACAG CACAGGAGTT GGGGGGAAGA ACAAGCCAAG  
 22501 TTCCAGATAT GGAGGTCAAGC ATGATTGTTG GCTTGAGCA GCCCACCCAGC  
 22551 TAAGGCAAAA AAGGAAACTC ATTGCATAGC TCTTACCTAT GGAAAAAGAA  
 22601 GAAATCTACT GGGGGCAGAT GGTCTTGTG GATTTTGCTG TTTTCTTTA  
 22651 TCTCCTTCC CAGCATTGTA TTCTGAGATA TTTCTCAATT TGGCTCCAA  
 22701 ATAAGCTTA TTGAGTGTGTT TAATGGTTA CTGTTTTTTT TAAAAATGGC  
 22751 TTTAACATAT AAAAGTACAA CTTATGGATC CTTTTTGTGTT GTGGTCGTGA  
 22801 CTTACTGATA ATATAATCCA AAATACATTT TTTATTTTGT ATTTATTAT  
 22851 TTATTTTGA GACGGAGTCT CAGTCTCTG CCCATGCTGG AGTATAGTGG  
 22901 TGTGATATTG GCTCACTGCA CCCTCCGCCT CCTGGATTCA AGCGATGCTC  
 22951 CTGCCTCAGC CTCCCTGAGTA GCTGAGACTA CAAACGTACG CCACCATGCC  
 23001 TGGCTAGTTT TTATACAAAAA TACGTTTTT AAAAACAAT TTTTTTTTG  
 23051 GAGGTCGGGG GACTGTCGCC CATTCTGTTG CCCAAACTGG AGTGCAGTGG  
 23101 TGCAATCTTG GCTCACTGCA ACCTCTGCCT CCCAGGTTCA AGCGATTCTT  
 23151 GTACTCAGCC TCCTGAGTAG CTGGAATTAT AGGTGTGTGC CATCATGCCA  
 23201 AGCTAATTTT TGTATTTTA GTAGAGATGA AGTTTCGCCA TGTTGGCGAG  
 23251 GCTAGTCTCA GACTCCTGGC CTCAAGTGTGACCT CAGCCTCCCA  
 23301 AAGTAGAAAA TCTTCTTGA AAATAAAATT CCAAATCTCA AAAGGCCCTA  
 23351 TATAATTTG GTGTTGGAAA TTTACTGTC AATGAAAATG ACTATTTACA  
 23401 CAAATTATAA GCTTCCATAT TAATATATAT GTGTGTGAAC CTGAAATTCA  
 23451 AATTTTTATAA TATTGTTTAT GAAAGGTACA GCCTCTGAGA TTCATCAGAT  
 23501 GGTATTTACC TTTAGGGCAT ATCTAAAAAT AAAATACAGT ACATGAAATC  
 23551 CAGTGTCTTA ATCCAGTGTAT CTTAAACTT TTTGCTCTCA GATCCCCTTT  
 23601 AAACCTTAA AAGATATTGA AGAGCTCCAA GGAGGCTTIG TTTACGTGGT  
 23651 TTTTATCAAT GGATATTTAC CATATTAGAC ACTGAAACTG AGGATTTAA  
 23701 AAAAATAA TTCATTTAA AATAACAGTA ACAAAACCCA TTACATGTTG  
 23751 ACATAAAATAA CATTTTTACG AAACATATT TTCAAAATT AGTGAGAGAA  
 23801 TGACATTGTG CTACATTGT TATAAACTC ATTATTGTCT GGCTTAATAA  
 23851 AACACTGCTG GATTCTCATA TCTGCTTTG NNNNNNNNNNN NNNNNNNNNNN  
 23901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 23951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24101 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24151 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24201 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24251 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24301 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24351 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24401 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24451 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24501 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24551 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24601 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24651 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24701 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24751 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN  
 24851 NNNNNNNNNNN NNNNNNNNAAA TATTGATTCA CTGATTATG TGGATCTTTT  
 24901 AAATGTTGAC ACTTATATAA TATAATACAA TATTTTAAA ATCACATTG  
 24951 TTAATTTAC CTTGATCTA TTCAGAAAAG ACTCTAAGTA TTGGGAACCT  
 25001 ATCATCCTCA CAGTGTAGA TACAAGTTTC CTAATTCTCT GATTTTACT  
 25051 GGAGAGCTCA AATTCTATCA TTGGAAACAA ATACACATT ATTAACTTA  
 25101 AAAATGACAG GATTACTTGG TTTCATATT GAGAAAATAC CTGTCAAATT  
 25151 CCCAAGTCTG GAAAACCATG GTTGTGATGTC ACTCTTCAA GTAAAAATGG

25201 CATTCCATGT AAGAAGTGTC TAGTTTATTA TGCAACTCAA ATAAATTACG  
 25251 CAAGTGCTT TCTTTAGGAC ATAACCTCAT ACATACTTCC ACAAGCAGCA  
 25301 GATGTGTGTA GTTATGCATA GTTCCTTATG CATGGTTCTT ATTTCATCAC  
 25351 ACAAAATATT AAAAAGACTC AGTGATTGAG ACCTAGCAGT TTTTACTGCT  
 25401 TCATCAAAGA TGCTCTTATT TGAAACTGGC ATAATATGAT TTATTTATTT  
 25451 GATTTACTG GGAAGCATGG CAGTCAGAA TGTAAATGACT GCCAGTACAT  
 25501 TTGAGTGCCA CTGCTTGATT TTTGCTATGG AGTCAGCAAT TTTGCCACTG  
 25551 GTTTGCATT TTCAGTAAAAA ATGTCAACAC AGTGAAGAAAG GCACATAATG  
 25601 TCTTGTATTA TTTGTAAAC AGTTTATCT TGAGACCCCC TTGAAAAGGT  
 25651 CTCGGGGATC CTCCAAGGTG CCAGTAGACC GTACTTTGAA AATCACTATT  
 25701 TTAATCCAAA GTGCCTAGAT CAGACACACT ATAATCCTG TGTCTGTAT  
 25751 GATCATTAGG TAAATACATT TGACTTAGA AGTATACATT CAGAGACATT  
 25801 AACAGTATTC AGGTTGGGAT TTAAGTATAT TTTAAAGTGT GGTACCTAGA  
 25851 GAGTATCCAT GACACTATGT TCATAAAATT TTAGAGAAAA CTGAGATCAA  
 25901 AGGAAACCAA AACAGGCTGG TCATAGTGGC TCATGCCTGT AATCCCAGTG  
 25951 CTTTGAAGG TTGAGGCAGA GGATCGCTGG ATCCCAGGAG TTTGAGACCA  
 26001 GCCTGGCAA ATATGGGAGAC TATCTCTACA CAACAAAACA AAAATTAGCT  
 26051 GGGTATAGTG TCTTGCCTGC ATAGTCCTAG CTACTCGGAA AGCTGAGGTG  
 26101 GGAGGATCCC TTGAGCCTGG AAGTTCTAAG TTACAGTGAA TTATGATGC  
 26151 ACCACTGCC CTCACCTGG GTGAAACAGC AAGACCCCTGT CACCCCTCAA  
 26201 AACAAACAAA AAACACTTTT TTCTCTGAGT ATGTAATATGG TTAGTGTACA  
 26251 GTCCTTGAAC ACATTGCAAA TAGTATAGCA ATATATGAAG TAGCCAGTAT  
 26301 GTGCTCTAGC TAATTTTATC AATCATCTCT TCCTAGACCA ATCAAATATT  
 26351 TTTCAATATT TTGATCCATG CTTATATGAA CAAGATTTTT TAAAGCTGGA  
 26401 AAATTCACA CATTATATA CTTACTATTG TTCTTAAAT TAATTTTTT  
 26451 TTTTTTTTTT TAAGCAGAGT CTTGCTCTT TGCCCAGGCT GAAGTTCACT  
 26501 GGGCGATCT CGACTCCCTG CAACCTCTGC CTTCCAGGCT CAAGCAGCTC  
 26551 TCGTGTCTCA GCACCCCAAG TAACTGGGAT TACAGGCATA CGCCACCACA  
 26601 CTGGCTAATT TTTGTAGTTT AAGTAGAGAT GTGGTTTCGC CATGTTGCC  
 26651 AGGCTGGTCT CAAACTCCCG GCCTCAAGTG ATCCACCTGT CTCAGCCTCC  
 26701 CAAAATGTTG GGATTACAGG TGGGAGCCAC TGCGCCCGGC CTACATTAAA  
 26751 TTTTAAAGCC TTTCTATGTC AGTGCATATA CCCAACCTAA TTCTTTTTT  
 26801 CCGTGAACCTT TTTGTGTTATG CTTGTAGCCT TCCTACCCCA GATTATTCG  
 26851 AAGCAAATTG TCATTCTGTA ATTTCAAATA TTACTATTTC AGTATTTAC  
 26901 AAAATGGTTG CAGTTTAATTG GTTGTCTCTT TTTTATTTAT TAGCTTGAT  
 26951 ATTTCTATAG AGAGTTTACCC ACACATCAAC CATTGGATT ACCTGAAGTA  
 27001 AGGGTGGTAC AGGAAAGGGAA GAAATCTTGA AATACTAGGT TCCTTAGCAT  
 27051 CCTCAAAGTT GACCAATGAG ATTNTTGTG TGTTGGTTG TTTTTCTG  
 27101 TGTCTCTGG ACTCATGGAT TTAAGTATAT TTGTTGGTTA ATCATCACTG  
 27151 TTATTATTCT TATTGATGTT CATGTTATT TAGATTAGTG GGAGCTTTT  
 27201 TAGTTGCTA TCTGTGTCTC TCGTCATGTC CTTAGATAAT CCTAATCCTA  
 27251 ATCCTGATTC ATCGTAGACA TTTCCCGAG CAAACCTGGA ATCAGCCATT  
 27301 TCTCAAGGAG CTCTCTGATT CCATTGAAGG AAAATATAAT ATAGGTACAA  
 27351 TCTAGGCACT AGGTGATACT TGTTACTTCT GGGTTGGCTA TTGTTCTAG  
 27401 CCTCCTAAAGT TTATATGACT GTACTAATTG GAATTCTAA CTATGGGACT  
 27451 AAACCTCTAA TTCTTAAATC TGCAATTCTCCT TTAAGTCATG CCAAAATCT  
 27501 GAACATCACA AACATAGTC TTTCGTTAC CCCACAATAC ACACATACAA  
 27551 CATTGTCAGT ATAACAGTAC CAACACCAC TCCAACAATA TGCCTACTGA  
 27601 AAAATTTAG GTAATCTGTC TCCAGCCTCC CAGGTAGCTG GGACTGCAGG  
 27651 TGCACACCAC CATGCCTGGC TAATTTTTT TTTTTTTTTT TTTTTTAAGA  
 27701 GACTGGGTCC TTGCTATGTT ACTCAGGCTG GTCTGAAATT TCTGGCCTCT  
 27751 AACAGTCCTC CTGCCCTTGC CTTCCAAAGT GCAGAGATTA CAGACCTGAG  
 27801 CCACCACGTC TGGCCTATCC TTTATTATT CCACCAAAGT TATTTATACA  
 27851 AATTACTTTG TTGTAAGTC CCTTGGAAATA GTTCTTCTG TGGCATTATG  
 27901 TTACCAAGTTA GATGCACCTT TGATTCTTT AACTTTACTT CAATTTTAA  
 27951 GGTTTGCTTT TTAGATTTAG TTTTGTTTA TTATACATAT ATGAAGTATT  
 28001 TCCACGGTTC CAAAGTTAAA TGAACAAAAC AGGCATGTTA AAAGAAGTCT  
 28051 AGTTTCTATC TCTGTCCCCT CCAACCCATT GTCTTCTTCC CCTTATAAGT  
 28101 AATAATTTCAC ATTNTTAACT TGTTGGTTAT CTTCTGATT TTAAAAAATAT  
 28151 AAGCATAAAAT ATTATATATTCT CTGTCCTTTA GCATGCTTT AGCCATCTG  
 28201 CTTTTTCCT GTATAATGCT AAATATATCT CATTCTTTTT AATTGCTGCA  
 28251 GAATTTCTCA TTACATAGGT ATACTGCAAT TTATTTATCT GATGCTATGT  
 28301 TGATGAACAT TTAAATGATT TCCAGATTTT AGGAACGGTG ATGATTGAAC

28351 TCTCTGTACA TATATCTTT TTACTGGTA CACTCCATCA AGCAACTACT  
 28401 TAAGTGAUTG ACTATGATGC TGTGCAAGCA GTTATATAAA GAAAACAGCA  
 28451 GTGACTCAGC CTGAAAACGG CTTAATATTAA TCATGTTTTC TTACACATTA  
 28501 TTTTATTGAA GGAAAAGCAA CATGGAGTTT AGTGATTATT TTTGAAAGAA  
 28551 ATAACCTATT TCTAATTCTA AAGAATGGTT ANNNNNNNNN NNNNNNNNNN  
 28601 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28651 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28701 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28751 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28851 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 28951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 29001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 29051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNN  
 29101 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNGAGTCTAG  
 29151 CTCTGTCAACC CAGGCTGGAG TGCACTGGCA CGATCTCTGC TCACTGCCAC  
 29201 CTCCGCCTCC CGGTTCAAG TGATTCTCCT GCCTCAGCTT CCCAAGTAGC  
 29251 TGGGATTACA GGCCTTCGCC ACCACACCCA GCTAATTCT GTATTTTAG  
 29301 TAGAGAAGGG GTTCACTGT GTGGCCAGA CTGGTCTTGA ACTTCTGACC  
 29351 TCGTGTCCA CCTGCTTCGG ACTCCAAAG TGCTGGGATT ACAAGCGTGA  
 29401 GCCACCACAC CTGGCCAAAAA ATATGGTTT CTAAAGCAAC AGTCCTAGTA  
 29451 CAACAGAAGA GAGGTGTTGA CTAGTTAGGG ATTAGGTTT AGAAGTACAT  
 29501 TCTTAGTAAG AGAGGTGAGA CTTACCTTCT TGTGTTTTAG TATAGTGAGA  
 29551 TCTGGATCAA ATCTATTACT CTTATTAAATC TCCTAACTTC CTACACTATA  
 29601 TCCAGTAGAG GACACTTTG CCTTACACAG TAAAGAAAGA GCCTCTGGAC  
 29651 TCTACCAATG GGATCGGAGC TCTCCAAACC TGATATTAA AAGGCCTATA  
 29701 AGTTTGGGG GGTCCCTTTG TCCACATGAT TATTCTGTAA TACATTGTAT  
 29751 TTATGGACAT GGTATTATTA TACACAGATC CTGTCTTTA AAGAACATTA  
 29801 TAATCCACTT AACTGCTAGG ACCAGAGAAAT GACCGATAAT TCAAACCATTA  
 29851 TTGCTTACA GAAGACATAT ATAAAAGATG GTATGTGTA CCAATTGAGG  
 29901 TTCAAATTG ATTCAATTAA AAACAATCTA GGGCAGATT TATATAGTTT  
 29951 GTGGACCCCTT TGCACTAAAC TCTCAAGGTT CTTATTAAAA TGCACTCTT  
 30001 GGCTGGGCAC GTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGCCAA  
 30051 GGCAAGGTAGA TCATTTGAGC TCAGAAGTTC AAGACCAAGTC TGGCCAACAT  
 30101 AGCGAGGCC AGTCTCATGG AAAGAAAAAA AATTTTTAA TAAAAAAATAA  
 30151 AAGCAGATCT TGGTAAAGA CATGTAGTCT GGTACAGG TATTAACAAAC  
 30201 TGTCTGTAAT GTAGTGAATT TGCTCCAGAC TTACCTTTTC CATTATTAG  
 30251 TTCTGAAATT ACTGTTCTAT GTATGGAAA TGAGAAAAAT TGCTAGATTC  
 30301 TAGAACTGTG GCTTCTATTIC ATAGTTGGAA AAATGAAGCA TAAACATTT  
 30351 TAATTCAGA TCAACAGCAA AAAGAAAGAA TCAGCATGGG AAATGACAAA  
 30401 AAGTTGTAT GATGCGTGGT CAGGATGGCT AGTAGTAACA CTAACAGGAT  
 30451 TGGCATCAGG TAAAGAAAAT TTTCAAGCA ATCCTTTTT AGTTAACAGA  
 30501 AGTATAAACT GTTCTCCCT CCTTCCCTCA ATTTTTTTTC AGGTACCAT  
 30551 GGATTTAAA AAGCATTGTT TTCTCTTCTT CAAAAAATCT CCTTAAATAT  
 30601 AAGACTAGGA GGCAGAGGGT TCCAAGTCTA GTCTGGCTC TATCACTTA  
 30651 CGTGTATTC CAGCTTGGTT GATCTTCTG GACTCAGTTT CTATATCTGT  
 30701 AAAATAAGTG GTTGGATCA GATGATCAAT AAAGTATCTT TTGATATTAA  
 30751 CATCGTAATA AATAGCTAAT ATTTCTGAG TGCTTCCTAT GTNNNNNNNN  
 30801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNTGAAAGAT  
 30851 TATGTCAGA AGACCATAAA AATTAAAATT TTTGTGGAGA ATAAAGTACT  
 30901 GATAATTCTA ATTGGCATGC ATAGTAATT TATGGCCTCT GTGTATGTAA  
 30951 CCCACTGATC TCTTCTGTAA AGAAGGACCC AGATTTGACC ATAAATTGTT  
 31001 GTATTTTTA TATTCTCACA ATAAAATAAT CTTGATATAT GGTTTTCTGT  
 31051 AATTAAGAA AATATTATTC CTATGAGTTT CAATAATTAT TTCTAATGGA  
 31101 CATTAAATTG TAATGAAATT GACATCATTG ATAAGTCTGT TAATTAAGTT  
 31151 ATCGATTGAA ATTGAGATTG GTGAACCTCC TGCCAAGTAG CTGTCCTTGT  
 31201 AAGATATTGTT AGTATCTTTT AAACATTGTT TTTCAGATCA CAATTAATT  
 31251 GAATGATGTA ACTTTTAAA ATTCCAAACAA AAAATAGCAC TTTTATTGTA  
 31301 AAAATAAACT CTTTACAGT TATAACTAAA ATTGAAAAAT CTTAAATTAA  
 31351 TATGATGTTA ATAAATGACC CTTTATTAG GAGTCTCCTG CTTTCTACTT  
 31401 GCCTTTAAC TAGATTGTT TCGACTCCC AAAAATTGAC TTAATTTTT  
 31451 TACCATCTCC AACATGTTT TATAGGGCA CTGGCCGGAT TAATAGACAT

31501 TGCTGCCGAT TGGATGACTG ACCTAAAGGA GGGCATTGCG CTTAGTGCCT  
31551 TGTGGTACAA CCACGAACAG TGCTGTTGGG GATCTAATGA AACAAACATT  
31601 GAAGAGAGGG ATAAATGTCC ACAGTGGAAA ACATGGGCAG AATTAATCAT  
31651 AGGTCAAGCA GAGGTAAGTC TTGCTTGTC TCAAGATGAA TTAATAATTG  
31701 ATATAGCAA ATGTTTCCAA TTCATTAAAT TATAAGACTA ATCACATATT  
31751 AGATGATTAC ATACACATCA AATGGATCCA CCCTCAACAC ATTGCAGCAA  
31801 GAAAGAATTA AGTGAATAT TGTTCAAGT AGCTTTTTA TTAGTTAACT  
31851 GCATAGTCAT ATAACAAATC CTCTGGATTG TGGTCAAAT ATATTTGAGC  
31901 TGTAGTAGAA AAGAAGTGT AGTTATTGCA GTAAGATCTG TGTAAAGTTA  
31951 CTAAGAAGTC AAGTTATTAA AACTAATATA TTACTAAGAT TGGGAAGTTT  
32001 GAATTATGAA AGTATTATCA AATAATTAG TAAAATCAAC CTACGTAGAG  
32051 ATACATTGAA GATAATCAGA CATTTTATT TGTGGCATTAA CAGCATTAA  
32101 ATGATTGATT TACTATGATC TACAAAGAAC ATTTTAAAGAC TTAGGATGTT  
32151 ACATGTATAT TTTTACATG ATGACATGGA TATATTTTT AAATTTGTT  
32201 TTAGCTGAAC TTTAGAGCTA AAAGGTATAC ATTTGCGGTA AGATGAGTAG  
32251 TATGCTGTTT CTCACCTGGC TTAATTGAAT TGAGTTTAAT GATCTGGAAA  
32301 GTTGCAGCAG AATGAAATCT GAGTGGTGT GCAATTGTT TCCACTGTT  
32351 CCAAAAAGTG GTTGTAGGC AGAGATTGAA GTATAGCTGA GATGTGTTGG  
32401 TAACAAGACT TTAGGGATTA GGAAAAAGAT TAAATGTGCT CAGGGTTCC  
32451 TGGTATATGT AGGCATTAAT TTTTGGACTC TACTTAAATA TTTTGTTCAT  
32501 ATAAAGTTT TATTATTGTG GAAATAAAC AGGAGACTTT TACACATTT  
32551 ACTGAAGTTT CTTTCTTTC TTTTTTTTT TTTTTTTTT TGGCCGGTGG  
32601 GATGGAGTCT CACTCTGTTG CCCAGGCTGG AGCGCAGTGG CACGATCTCG  
32651 GCTCCCTGCA ACCTCCGCTT CTGGGGTTA AGCGATTCTT CTACCTCAGC  
32701 CTCCCGAGTA GCTGGTATTA CAGGCCTGCG CCACCATGCC CAGCTAATT  
32751 TTGTATTTT AATAGCAACG GGGTTTCACC ACATTGGCCA AGCTAGTCTC  
32801 GAACCTCTGA CTCAGGTTA TCCACCCGCC TCAACCTCCC CAGTGCTGGG  
32851 ATTACAGGCG TGAGCCACCA TGCCTGGCCG TTTACTGAAAG TTTCTTATGA  
32901 CAAGCATTG CATTAGAGGT GCAATGAAA TAAATTCAT ACTCTCGAAC  
32951 TATTCTCTT TTAGGGTCTT GGTCTTATA TCATGAACTA CATAATGTAC  
33001 ATCTCTGGG CCTTGAGTT TGCCTTCTT GCAGTTTCCC TGGTAAAGGT  
33051 ATTTGCTCCA TATGCCCTGTG GCTCTGGAAT TCCAGAGGTA AGCCAAGTAA  
33101 TATTAGTGT CATTAAACAT TATTATGATG CTTATCTTT TGACCTTAGT  
33151 GATAATAAAA GTTGGCTTT CTGGAGGGAG GGGATAGTTT GTTCATAATA  
33201 TGAAAAAAA ATTTTTTTAA GTATAAGCTG ATGGTAGACA TCATTGAAAAA  
33251 ATATTGTTCC CCATAGTCAT TTGGTCATTT ACTGTGAAGG CTGATTTTT  
33301 TTTTCTCTCA CCACTAATT AACACATGAC TAGGCAAATT TTCAGACTAT  
33351 TTAGTTAACAC ATCAAGAGCC TGGAAAGAAGT ATCTTGAC GCTAATGTTCT  
33401 TTGACGGGTT AGTTGTTACT TTGCTGTTAT GACCCTGAAT TTTTTTTTT  
33451 TGAGACTGAG TCTTGTGCTG TCGCCCAGAC TGGAGTGCAG TGGCGCAATC  
33501 TCAGCTCACT GCAACCTCTG CGTCCCAGGC TCAAGCAATT CTTGTGCTC  
33551 AGCCTCCCTGA GGAGTTGCGA TTGCAGGCAC CTGTCACCAT GCCCTGCTAA  
33601 TTTTGTCATT TTTTGTGTTG TTTTTTTTT TTAGTAGAGA TGGGGTTCA  
33651 CCATGTTGGC CAGGCTGGTC TCAAACCTCCT AACCTCAAGT GATCACCCGC  
33701 CTCAGCCTCC CAAAGTGTG GGATTACAGG TGTGAGGCCAC CACACGTGGC  
33751 TATGACCTGTG ATTTGATTC ATTCACCTTT TATAATTACC TTTTGATTAG  
33801 ATAAGTTAAT TATTCTGAA TTTGGCCATT TTATGCTTTG AGAAAGTAGT  
33851 TAATCACAGT GGGTCAACAG TACAAACCTT TGGGTTTTAT TTTTCATCAC  
33901 AATAAAAGTAG AGTTATACAT AGGATTGATT GAACTTGATT TGAACCTTATC  
33951 TCTTCTCTT TATTGTTCTG GAGTTAAATA AGTTACCAAC TTTTCTCTAA  
34001 TACATTCTT TTTAAATGG AATTGTATTG ATCCCTTAAG TTTGTATTA  
34051 GAATATCTT CATAAAAAGC AATATCATGC AGTATATAAC AGTTGTTACT  
34101 CATTCTTGAT ACATAAAAAT CTATTGCACA TAATTACAGG ACCTCAGAGA  
34151 AAACATAATA TTCTTATTTC TAACATAATG GCCAAAATAT ATTTAAAATA  
34201 TTATGCTTAT TTTTACAACA GAAATATTCA AATTGCCCCCT TTTTTGGGT  
34251 ATGTAATTAT AATCCTTATA ATTAAGGTCT GTATTCAATT TAACATGCC  
34301 TGATATTTG ATTTGGCCT GAGATAGTGT TGCCCTCTCT CTTTCTTGG  
34351 GTAGAGAATT AGATTATAAT ATCAATTAT TATATGTAGC ATAATAGGCA  
34401 AGTTTCGAA AAATTAACGT TAAATTGTT TGAGACTGTC TAAAATTGTC  
34451 AAGGGTGTGTT TTGTCATAA ACAAGAAAA TAACTGGAT TCGTTACATT  
34501 CTCATGTTTC TAAAGGACA TTAAGCTGCC TTAATCTTTG CCTTGAGAT  
34551 TAAAACCTATT TAAAGTGGAT TCATCATCAG AGGTTACTTG GGAAAATGGA  
34601 CTTTAATGAT TAAAACCATC ACATTAGTCC TGGCTGTGGC ATCAGGTTG

34651 AGTTAGGAA AAGAAGGTCC CCTGGTACAT GTTGCTGTT GCTGCGAAA  
 34701 TATCTTTCC TACCTCTTC CAAAGTATAG CACAAACGAA GCTAAAAAAA  
 34751 GGGAGGTAAG TGTCTTTGT AGTTAATTG ACTGAAAAAT ATATATTATA  
 34801 TAGTATTAT TTAAGTAAAG AATTCTTAG TGAAAAATA ATAAATTCTG  
 34851 TATTCAAGATA AAAAATTTG AGATTGTGC TTCTGTTTT CCTGAATAAT  
 34901 CTATAACATC TTTCTAGAAT CCATCCCCAG TGCTGCTCAG TTCGTCTAC  
 34951 ATTTAGAGA AGCTTCTAGAT AGACAGCTGG TGCCATTGG GTTCAGCTG  
 35001 CATTCAAGA AGATCTCCCT GTTATCACTT TACCTTACAT CTTTCCTCTT  
 35051 CTGAAGTGT TTCTAAGCT AGCTTGTGTT TTCACTCTTA CTTTCAACAT  
 35101 TAAGAGGTTG GGAAATCTTA ATAGCTATGT TTTCTCCTG GAGGCAGTGT  
 35151 CTGGGCCAG TGTAAGTGGT GTGTGATATG AAAAATGCTA TCCAGTGCTA  
 35201 TGGGAAGTT CTGAGGGCT TTAGAACGTC TTGAAGTTA AATCAGAAAT  
 35251 TCACATTAAA GAGATTACAG GAAATCCTT TCATTTGATT GTTTAAGGCA  
 35301 ATTTCTTTA CCATTTCTT AGGCCAGCCT GAGATCTTCT ACAAGACCTT  
 35351 GAAACCTTAT ATATATTATG GATTTCTCT GATGTTCCA TATTGCTCTG  
 35401 GGCATTTCC TGAATCCTT ATATTAGCTC TAGACTTTGG GAGCCCAGTC  
 35451 CCTTCCTATT TTCCAAATCT AAATCTACAG CCCTAGATGG TACAGAGATC  
 35501 TTTGAGTTT TAAGATATGA TTTTTGAAA AACATCTCAT TAAATACTGG  
 35551 CAGAACCTT TCATCTTGT GAGTTTTTA ATGTAATGTA ACCAAAAAAG  
 35601 TAGAATATTT TATCAAACCTG TTTAATCTTC AATTGAAATA ATTCTAGTAC  
 35651 ATTTTAATGT TCGCATTAAA ATATTGCTCT TGCAATTGGAC GTAGATATCC  
 35701 CAAAAGTGGA ATACTTCAGA TTGCTGTAGT TTCATCTCTG AATAATTGTG  
 35751 ATTCCAGTAC TTTATAACAA AAATAGCTAG CATTATTGAT TACTTTCTGT  
 35801 GTATCTGGTA CTGTGGCAGA TACTTTACTT GGATTTTAAT ACTTAATTT  
 35851 ACAGTAATT AGTAATATGG CCCTGTTATC CTCATTTAGT GATTAGTAAA  
 35901 CTAGGGCTGA AAACAGCTAA CTAACCTGCC CGAGACTACA TACCTAGTAA  
 35951 GTGGTGGAAC GTAGGTTAAA ATTCACTTTT CTTGACTTC AAAGTCTGTG  
 36001 GTCTTACCTA CTTACATTAC TGCCCTTACG ACTATGTGGG TATATATTG  
 36051 TGTGTGTTCA AAACAAACTC AAAACCATCC TGAGCGTAG CAAGTTAGTG  
 36101 GCTAAGATGA AGCTAGAGCA TTTGCCCTT CAATTCAATT CCATTACTTT  
 36151 CTGTTGTAC TTTATATTGT TTGGTAAGAC TTTTACTTAT TCTAAGTCA  
 36201 AAAAATGTA TTTATTAGAT GTTTGAGAAA TTAAGTTTAC CTAAATTAA  
 36251 ATGTTCATAC TGTAGTGTAGT AGTTAATGTT TAATACGTT TGATTCTGTC  
 36301 ACCTTAGTGT ATATATAAAAT GGCAAGAATT CACGGTTAGT TGAAAGCATT  
 36351 AAGGCCCCAT AGTTTGTGT AGACAAGAGG GGAGAGCGTT GATATTTTA  
 36401 AATTAAATGTC TTCTTAGATA CGTATGAAAT GGATTTAAAC ATGTATATGA  
 36451 GTTATAGATA CCTAGGTGT AGTTTGGTT TAAATTCAAGG ATCAGGACAT  
 36501 TCAAATAAAAT ATGTTGCTT TCCCTTTAGT GGAGGAAAAA AAAAAGAAC  
 36551 TAAATTGCT CCCTTCCCT CCCAAATAAC CAGAGTCTAC ATTTTAATGC  
 36601 CAACAAATTG ATTAAACAA ATATTATTT ATTTTAATT CACCAAACCTT  
 36651 TTATAAAGTA TTTACTGGTG CCAGGCACTG TTCTAAAGCA CTCTGTATAT  
 36701 ATTTACTCAG TCCCTTAAGAG CTAAGTAATA TTATCACGTT TCCATTCTAG  
 36751 AGAAAACGTGA GGCACATATA GGTTAGGTTA TCTACCCATA GCCATACAGC  
 36801 TAGTAAGTAG CAGAGCCATG ATTCAACAC AGCAGCCTGA CTATGGAGTT  
 36851 CATGATCTTA ACCATTACCA GCTTAATT TATTATTTAT AATTCTCTT  
 36901 CTGAAATGT ACAATTGAC CATTGAAAGA AATACTTTAG GTAGCTTGG  
 36951 ATATTTGCTG TATTAAAGTA GTGAAAGTAA TACAGACACT TGGCTGGCG  
 37001 CGGTGGCTCA CGCCTATAAT CCCAGCATTT TGGTAGGTT AGGCAGGGCAG  
 37051 ATCACCTAAG GTCAGGAATT CGAGACCACTG GTTCCAACA TGGTGAACC  
 37101 CCGTCTCTAC TAAAAATACA AAAATTAGCC GGGCGTGGTG GCAGGCGCCT  
 37151 GTAATCCCCA GCTACTCGGG AGGCTGAGGC AGGAGAACCA CTTGAACCCA  
 37201 GGAGGTGGAG GTTGCAGTGA GCTGAGACGA CGCCATTGCA CTCCAGCCTG  
 37251 AGAAACAAGA GAGAAACTCT GTCTAAAAAA AAATAAAAGGA ATACAGACTC  
 37301 TTAGAAAAAT AATTACAAAT AAAACCTAG TGAAATTATA GGTATAGTTA  
 37351 GGTATAGTTG GCTTACAGGT GGGAAAGTAGA CCATTACCAA CTGATAGACT  
 37401 GGGGAGCTGG AGAGAGGACA CGGAAGAGTG TCCTTGGATT TTTCNNNNNN  
 37451 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN  
 37501 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN  
 37551 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN  
 37601 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNN  
 37651 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NAAAATTGTC TATATTCTT  
 37701 GCCTCCTCCT CTTTACACCC TATTCACTT AGTATATCTG GCAAAATTT  
 37751 TTTTAACCTG AATGGTAAAT GCATGACTGA CTTTCAATT AAAGGCCAGGA

37801 GAAAGAAAACA AATCTTAATA GAAGAAATGA ATAGTTACCC TTTGCTTAGG  
37851 GAGCAAGGAA ACATGCAAGT TAAATTCAAGA AAATCCATTG GGAAAATTCA  
37901 AGTAACATGA AGAATTTTTA TTTGGTATGT TTGAATTCT ATGAAATTAT  
37951 GAAATAAGCC ATATCCTCTT TCTAGGTGCT ATCAGCTGCC TCAGCTGCAG  
38001 GGGTTTCTGT AGCTTTGGT GCACCAATTG GAGGAGTTCT TTTTAGCCTG  
38051 GAAGAGGTAG GTGAAAAGAA TACAACAAATT AAAATTATAT ATAATTACCA  
38101 TTACAAATAT ATTCACACA TTTCAGTTT GTAGGTGATG TAATAGGTAG  
38151 AGACTTTGTT TCACAAATTAA TTTTCTAAA GTTGTTCCT ACTCATTCTT  
38201 AATAAAAAGT AAATGTTATT CATGCTCCAT ACCTGGAGGA AACTTTTAA  
38251 AAATTATTAA ATGTATGAAT GTTAGTAAATT ATTAAAATC TAACTTTGTT  
38301 GACATATTAA AAAGTAAGAA GATGTGAATT TGACTTAATA GAGGACATGT  
38351 GAAACAATCT ATTTCCATTG GCTAAATTCT GTATTTTTAG TAGAGATGGA  
38401 ATTCACCAT GTTGGCCAGG CTGTTTTT GTGGGGTTTT TTTGTTTGT  
38451 TTTGTTTGT TTTGTTTTT GAGACGGAGT TTCACTCTTG TTGCCAGGC  
38501 TGGAGTCAA TGGCGCGATC TTGGCTCACT GCAACCTCCG CCTCCAGGGT  
38551 TCAAGTGATT CTCCGCCTC AGCCTCCCAA GTAGTTTTG TTTAAAAAAAT  
38601 TTTAATCAAT TCCTATGTT AGTTTAAAG TTTTCCCAT GTGATTATTT  
38651 CTGATACAGT TAGTGATGTT AAAGAAAATA ATTTTAGTGA CTTCAGTGGA  
38701 TTATTTGTT TTTGTTTCT TAATAGGTGT TTAAGACTTT TCTTTTACA  
38751 TAAAAATGTA ACCAGGAATT TTTTTTAAT TTTTTTGAC AAATAATAAT  
38801 TGTTTTGTT TATGGGGTAT AATGTGATGT GTCTATACAT GTATACATTG  
38851 CGGAATAATC AAATCAGAGT GATTAGCAA TCCCTCAAAT ATTTATTATG  
38901 TCCTTGTGGT GGTGAGAACAA TTTAAAATCC TCTTTAGCT ATTTTGAAT  
38951 ATATAATACA TATTATTAAC TGTGGTCATC TTACTGTGCA ATAGAACACC  
39001 AGAACTTATT CCTCCTCTGT AAGTCATAC CCGTTGACTA ATGTCCTCCC  
39051 TTTCCCTGTT CACCTCCCCA ACCCCTAGCC TCTGGTAACC CCTATTCTAC  
39101 TCTCTACTTC TATGAATTAA ACTCTTTAG TTCAAGATGT TTTTAAATGT  
39151 ACTTTTTCT TTAGTTGTT TGTATTCTT TTTTTTTTTT AATGTAGAAG  
39201 AGGCAAATTA AATGCATTAT AAGTTAACAG GAGTTGGTGA TGGTACATT  
39251 ATTTTAACCT ACCATGATTG AATTGAATGT GAAACTCATT TTGAATATAA  
39301 AACAGCACTA GGTATTCTAT TAGTATTAT TAGACATTAA TGATCAATTG  
39351 ATACTGTCAA TTGTAATGA TGATCACCAT CTCCAAAAT AATAATAACA  
39401 TCAATTTC TTATTACAGT AAAATCATT ACATGTAAT TCTAACTACA  
39451 GCAAAATTAA GAGCTAGGAT ATTTACATT CAAGTTATAA TATATCAGAA  
39501 ACATCTTATA AAATTATAGC ATTAATTTTT CTTTCCCTTT TCTTTTTTTT  
39551 AGGTTAGCTA TTATTTCCCT CTCAAAACTT TATGGAGATC ATTTTTTGCT  
39601 GCTTTAGTGG CTGCATTGTT TTTGAGGTCA ATCAATCCAT TTGGTAACAG  
39651 CCGCTGGTC CTTTTTTATG TGGAGTATCA TACACCATGG TACCTTTTG  
39701 AACTGTTTCC TTTTATTCTT CTAGGGTAT TTGGAGGGCT TTGGGGAGCC  
39751 TTTTCATTA GGGCAAATAT TGCCCTGGTGT CGTCGACGCA AGTCCACGAA  
39801 ATTTGGAAAG TATCCCGTTC TGGAAAGTCAT TATTGTTGCA GCCATTACTG  
39851 CTGTGATAGC CTTCCCTAAAT CCATACACTA GGCTAAACAC CAGTGAACGT  
39901 ATCAAAGAGC TTTTACAGA CTGTGGTCCC CTGGAATCCT CTTCTCTTG  
39951 TGAATACAGA AATGACATGA ATGCCAGTAA AATTGTCGAT GACATTCTG  
40001 ATCGTCCAGC AGGCATTGGA GTATATTCAAG CTATATGGCA GTTATGCCTG  
40051 GCACTCATAT TTAAAATCAT AATGACAGTA TTCACTTTG GCATCAAGGT  
40101 AAGTGCTAAT GTGAGGTGAT ATTTGGTAA TTTTGGCATG TTCAAAACTT  
40151 ATATGTGAA TGAGAGAGGT TGTGTTTCA TAAATGACTG AAAAAAGTAC  
40201 TTATCTTTG AGTTTAATT TAAGTAATGA AAAAGATAAT TCCTTAGCAT  
40251 ATATTGTTGA CCATGTTATC TGTGCTATT TAACAAATTA CCCCCCAAAA  
40301 CTTAGCAGCT TAAGGTAACT ACTTATTCTT TTCTTGATAT TGAGTCAACG  
40351 ACTTGGGAAG GGCTCAACTG GGCAATTCTT GCTGTGGTC TTTCATATAG  
40401 TTGTTATTAG ACATGGCGAG GGCTAATCAT CTCAAAGCTT CTTTTTTTCG  
40451 TTTCTTTTTT AAAAAACTGT TTTTGTGGAT ACACAGTAGC TATATATAGT  
40501 TTTGGGGTAT ATGAAGTATT TTGATAGAGG CATGGAGTGCA ATAATAATCT  
40551 CAGGGTAAAT GGAGTATCCA TCACCTCAAG CATTATCCC TTGTGTTACA  
40601 AACAAATCCAA TTACACTCTT AATTATTTT AAGTGTACAA TTAAATTATT  
40651 GAATATAGTT CAAAGACTTC TTCATTCACT ACTAGCACCT AGGCTAAAAA  
40701 AATTCAAGACA CCTGGGCTCC TGGGATCAAT CACCGCATACT GTGTCTCTG  
40751 TGCTCACTCC CGCTGTCTCT CTCTCTTCT CTCGCTTCTT TTTCTCTC  
40801 TCTCTGTGGT TTTCTAGGGT GGTGGCCTCA GGGAAATTGGA TTTCTTATAT  
40851 TATAGCTCAG GATTCCCAAG AGGGCTGTTT TTAATGTCAG CAAAGAAGTC  
40901 TTGCAAGCGTG ACTTGTTTA TTCTATTCA TGAGGTAGTC ACAGAGGCC

40951 GACCACATTC AGAGGAGGGA CATACACTTG CTGGGACAAG TGTAAGAGAA  
 41001 TTCACTGATCA TGTTTTAAAA CCACCTTTAT TAGTTTCCCTA TTGCTGCTGT  
 41051 AATAAATTAC CACAACCTAA TGGCTTAAAA GCCACACAAA TTTAATATCT  
 41101 TACAGTTCTG CAAATCAAAA GTCTGAAACG GATCTCACTG TGCTAAAATT  
 41151 AAGGTGTTCG TAGGGCATTG TGGAGGCTGT AGGAGAGAGT CTTGTTTTT  
 41201 GCCTTTCTG GCTATTTAAA GCTGCCAGCA TTCCCTTGGCT CCTGGCTGTC  
 41251 TATTGTCATC TTCAAAGCCA GCAGTAGCTG GTCAAGTCTT TCTCTTGCT  
 41301 CATCACCCCTG ACCCAAAACTC TGCTAAATCT CCCTTCCACA TTTGAAAAAC  
 41351 CTTTGTGATT ACTTTAGGCC CACGCAGATA AATCAGAAAA TAATCTCCTT  
 41401 TTTCAAGGTC AGTTGCTTCG AAACTTCTT TCTGCCACCT TGATTCCCTC  
 41451 TTGCATGCA ACCTAATGTA ATCACAGGT CTGGGAATTAGTTAGTATGGAC  
 41501 ATCTTGATG AGCCATTATT CTGCCTCATC CCAGTATAGG GTATTAGCTT  
 41551 GAAAGGACAC TGCAGACTCA GTTAAATTAC TAGATCTATA AATACATGCC  
 41601 TTTTCCATC AAGAAATTAA GGCAGCTGGG TCTTATGCCG TGGGACATTG  
 41651 CTTCTTTGG ATTTATAAAA TAACAAAATT TGTTGATTAA TGGTCTATCA  
 41701 GTAAATATAA TTTCTTATGT GACTATCAGT GATATATATG GGGAAAGCACA  
 41751 TATCAGCTTA TTCTTGTCT TTAAATTACT ACCCTGTAC TTCATGTAAT  
 41801 AGTATTTGCT AGTGTGATG TGCTTTTACA GATGTAAATT AATGTGGAAT  
 41851 AACAGCTTTG TTTCTACAAA ATTAGAGTGG TTTTAGTTTG TGAAATAAGG  
 41901 TCTCTTTCT CTTGTCCTAA GTCTGTAGTC CACTGAGTAT CTAGAGTAA  
 41951 ATAATAGAAA AGCCTGGCCA GGCAGCTGGG CTCACACCTG TAATCCCAGC  
 42001 TCTTGGGAG GCGGAGGCGG GCAGATCACA ATGTCAGGAG ATCGAGACCA  
 42051 TCCTGGCTAA TGCCTGAAA CCCCGTCTT ACTAAAAATA CAAAAATTAG  
 42101 CCAGGCCTGG TGGCAGGTGT CTGTAATCCC GGCTACTCGA GAGGCTGAGG  
 42151 CAAGAGAAC TCTAAACCC AGGAGGTGGA GGTGCAATG AGCCAAGATC  
 42201 ACACCCACTG CACTCCAGCC CAGGCAACAG GGCAGACAC TGTCTAAAA  
 42251 AATAATAATA AGAAGAAAAT AATAATAGTA ATAGAAAAGC CTAAACATTT  
 42301 TACCTTTTT TCTTAGGGAA TCAAGTTAAAGAGCTGTTA AAGCTCTTTT  
 42351 TCCTACAATA AGTAAGTGT GGGTAAATCC CAACTTTCTC ACAGTCAGTT  
 42401 GAACTACAAG AAGCTGGAGG CAATTGGCAG GCCTTGTGTT AGTCCCACCT  
 42451 TTGACTCAGC TCTGGCTGAA GGATCATACC TGCAAGAGA GTGTAACACA  
 42501 CACTTGATT TTTCTATTG TTTATCCTT TAATGATCCT AAGAGACTCA  
 42551 AGAGTACATG CCATCATTGT GTGTTGGCT CATTTCATAT TCAGAGGAGT  
 42601 TTATTAATCTC TTCACTGTTGTTGTTGTT TTTGAGACA  
 42651 GGATCTGCC TTTTGGCCA GACTAGAGGG CAGTGTGCA GTCTTGGCTC  
 42701 ACTGTAACCT CCACCTCCCCA GGTTCAAGCG ATTCTCCTGC CTCAGCCTCC  
 42751 CAAGTAGCTG GGATTACAGG TGTGGGCCAT CACACCCGGC TAATTTTGT  
 42801 GTTTTAGTA GAGATGTGAT TTGCCATGT TGCCAGGCT GGTCTGGAAC  
 42851 TCCTGACCTC AGGTGATCCT TTGGGAGGCC TTGGCCTCCC AGACTGCTAG  
 42901 GATTATAGGT GTGAGCCACT GAACCTGGCC TCTTCAGTA GTCTTTAAAT  
 42951 GATCTTGCTT ATGGTGCCTC TTATCCCTGT TTATTATCCT TATTAAATT  
 43001 AATCAATAAA TATTTTCTC TTTTTAATTG ATTCAATATAA ATAGACTTAC  
 43051 CTGAGAGATA TAGGTTCACTG TCAGAGCACC ACAATAAAAGT GAATATCATA  
 43101 ATAAAGCAAG TCACATAAAA GTCTTAGTTT CTTAGTGCAT ATAAAAGTTC  
 43151 TGTTTACACT ATGCTGTAGT CTTATGTGTA CAATAGCATT ATGTCCTTTA  
 43201 AAAAAGTAAT ACTTTAATT AAAAATACTT GATTGCTAAA AAATGCTAAT  
 43251 AGTAATCTGA GTCTTCAGTG AATTGTAATC TGTTTGCTT CTGTAGGGTC  
 43301 TTGCCTTGAT ATTGGTGGTT GCTAGAGGTA GGACTGGCTG TAGCAATTCT  
 43351 TAAAATAAGA TAACAGTGAATTTGCCGCA TTGATTGACA CTGCCTTCA  
 43401 TGAAAGATTT CTCTGTAGCA TGTGATGCTG TTTGATACCA TTTTACCTAC  
 43451 AGTAGACCTT CTTTCAAAA TTAGAGTCAT CCTCTCAAAC CCTGCTACTG  
 43501 CTTTATCAAC TAAGTTTAAG GAAAATTCAA AATCTTTGT CTTTTAAACA  
 43551 ATGTTCACAA CATCTTACCA AGGACTGGAT TCTACCTCAA GAAACCACTT  
 43601 TCTTTGCTCA TCCATAAGAA GTAACCTCTT ATACATTCAA GTTTTTAA  
 43651 TGAGATTCTA GCAATTCACTG CACATCTTAA GGCTACGCTT ATCATTCTAG  
 43701 TTCTCTTGCT ATTTCCACCA CTCTGTAGTT ACTTCTTCAA CTGAAGTCTT  
 43751 GAACCCCTCA GAGTCATTCA TGAGAGTTGG AATCAACTTC TTCCAAACTC  
 43801 CTGTTAATAT TGATATTTTG ACCTCCTCCC ATGAAACGTG AATGTTCTGG  
 43851 ATGGCATCTA GAATGGTGAC TACTTTTGA ACATTTCAA TTTTATTTGC  
 43901 CCGGATCAAT CAGAGAAGTT GTTATCAGTG GTGGGTTTCC AAGTTGTCAG  
 43951 GGGCGAACCA TACAGATCTT CAGCAACCTC AACTCTTGCC TTCTCAGAGG  
 44001 AAAGAATTCT ACGGAGGGAC ATAAGGCAGA AAAAGAGACT GAGGCAAGTT  
 44051 TTAGAGCAGG AGTGAAGTATT TATTATAAA AAGCTTCTAGA GTGGGAATGA

44101 AAAGAAATTA AAATACACTT GAAAGAGGGC CAAGTGGCA TCTTGGAAAGA  
 44151 CAAGTGCCCT ATTTGACCTT GGACTTAGGG TTTTATATGT TGGCATACTT  
 44201 CTGGCATCTT GCATCCCTAT TCCATTGATT CCTCTTTGG GGTGAGTGC  
 44251 CCACATGCTC AGTGGCCTGC TAGCACTTGG GAGGGGAGTG TGACAGTGT  
 44301 ATTTACTGGA GTTGTATGCA TGCTTACCTG AGGTGTTTGT TGCTTACAG  
 44351 CCAAATGTCC CTAGGAGGTC ATATTCAAA ACTCCATGAT TTTGCCTCA  
 44401 AATGTGCATG CTTGAGGCCA CTCACCCAAC TCCTGGGATC TTATCGAAA  
 44451 GCTGCCGATC GCTAGTTCA GGTGTTCTA TCTATTGGAA GATGGCCTT  
 44501 CCCTGATGCT GGCTGCAACC AATTATTACT TTAGAGAGAG AGCATGAGAG  
 44551 CTGCTCTCACC ATCATCACCT GATGGTGCC TGACATTCTC GGTGGGGTTG  
 44601 GGAGGATGCC TGTCTGCC TGCTCATGCC TGACTAGCTA CCTGCTGTAA  
 44651 CAAAAGTACT ATCTATGGTA GCTGTAGCCA TAGGAAATGC ATTTCTTCAG  
 44701 TAAAACCTAA AAGTCAAAAT TAGTCTTTAA AACAACATGA ATCTCCTTGT  
 44751 ACATCTCCAT CAGAGCTCTT GGAAGACCAAG GTGCATTATT AGTGATGAGT  
 44801 AATGTTTAA AAGGAATCTT TTTGTCAG GAGTAGGTCT CAACAGTGGG  
 44851 CTTAAAATAG TTAGTAAACC ATGCTGAAA CAGATATGCT GTTATCCAGG  
 44901 CTTTGTATT CCATTTATAG AGCACAGAGA GAGTAGATTG GCATAATTAA  
 44951 AGGATTACTT AAAAAAAAG TCTTTGATTA CTCTCAAAAA AAAGTCACGT  
 45001 CTCTCACTTT ATATCAACAG CTAAAAATGG CCAGGTATTG TGGCTCACGC  
 45051 CTGTAATCTC CATGCTTGG GAGGCCAAGG CAGAAGGATC ACTTGAGGTC  
 45101 AGGAGTTAGA GACTAACCTG GGCAACATAG TAAGACCCAT CTCTACAAAA  
 45151 AAAAAAAAG AAAAAAGAAA GCCAGGTGTG GTGGTGCACG CCTGTAGTCC  
 45201 CAGCTACTCA CGAGGCTGAG TCGGCAGGAT CACGCCAGC CAAGAGACGT  
 45251 GACTTCTGCT TTCAGTTGTA CACTTAGAGA CCATTGTAGG GTTCTTAGTT  
 45301 GGACTAATT CAATATCATT GGGTCTCAGG GAATAGGGAA GCCTGAGAAG  
 45351 AGGGAGAGAC AGGGGAACAG CCAGTTAGTG GAGCAGTCAG ACCACATACA  
 45401 ACACCTTATT AGTTCACTTT CTTCTATGGG CATGGTTCAT GGTGCAGTAA  
 45451 AACAACTGTA ACAGGAACAT CAAAGATCAT TAATCACAGA GCACTGTAAC  
 45501 ATATAATAAT AGTAAAATT TTCAAAAGTAT TGAGAGAATT AGCAAAATAT  
 45551 GATACAGAGA CACAAAGTGA CCACATGCTG TTGGAAAAGT AGTGTGATG  
 45601 GACTAGTTG ATGCAAGGAT GTCATAAACCC TCAATTGTTG AAAACTGCAA  
 45651 CATGTGTGAA GCACAGTAAC ACAAAAGCATA GTAAAACAAG ATATGTCGT  
 45701 ATATCAGTCA AAATATTGGG CAACTCTGAT AAGTTGTCC ACTTAACATT  
 45751 GTACCACTTA AGATGAATAG CATCTACCAT TTCCGTCATT TGTAATATA  
 45801 TAGGAGGACA TAATCACATA ATCTTGAAGT AAAAGACAGT GCTTAAACT  
 45851 GAATCAGTTA AGTTTATGA AAAATACCTC ATATTGTACT TTTAAAATA  
 45901 TATATTTTT AATTCATAA GCTTTGGGT TACAAGTGGT TTTGGTTACG  
 45951 TGGATGAATT CTATAATGGT GAAGTCTAAG ATTTTACTGC AACTGTCACC  
 46001 CAAGTAGTAT ATATTGTATC CAGCATATTG TCCTTTTTT TTTCTTTTTT  
 46051 TTTTTTCATT TCACCATGGA CTAATGAAAA TTTTGTTCAGG GACTGACATT  
 46101 AGGGCACCT TGAGCTACCT TGAGCTAAAG GAAATAACCC TTGAATTTTT  
 46151 TTCTGTTGG CCTAGAGAAAT GTGGTTGTT TTGTAACTGA ATTCAATGGGA  
 46201 TTGTTAAGGT ACAAGATTTT GCTTTAGTT TATTTGTACT AGGATTTG  
 46251 TATATTAATA CAATGTGAAA AGAATCAAAA GTGTTAGAAA TAAATGCATA  
 46301 GAATGTAAGT TTCAGGCATG TGAGTAGAGG ATCTCTGCTC CATAAAAGAGT  
 46351 TCTGTTGTTG TTATAGGTT CATCAGGCTT GTTCATCCCC AGCATGGCCA  
 46401 TTGGAGCGAT CGCAGGAAGG ATTGTGGGGA TTGCGGTGGA GCAGCTTGC  
 46451 TACTATCACC ACGACTGGTT TATCTTAAG GAGTGGTGTG AGGTGGGGC  
 46501 TGATTGCATT ACACCTGGCC TTTATGCCAT GTGGTTGTG GCTGCATGCT  
 46551 TAGGTAATAT GGCTGTGTCT GCCTGTGTGT GGATGTTTGC AAGTCTGAGA  
 46601 GAGCCAAGAG AAAGTGGGAC ACATTCTGC TTAATTGGT GGCAGGATTGG  
 46651 TTGAGTAAAG GAGGGTGCCTA GGAGGGAGATG TTTAACAGA TAAGAACAG  
 46701 TAGTACTATT AGGGTATTAT ACAGTACCGG TTTTCTGTCT TACAACATT  
 46751 GTTAATACAA GAATTTAATG GCATTAGCAT ATTGTAATAT AACTTAATAC  
 46801 ACTATGGCAG AAGCCATCTA AGTACAAACAT AAGCTTAATT TGAATCCTGA  
 46851 CCAAAGATGT CTTGATTCT TTCATCGTTA AGGATCTTGG CTTACCTATA  
 46901 ACAACTATAG CATAATACCT AAGATTAGCA TTGCAACAGA GTTTCAGAGT  
 46951 AGGTTTACTT TGGTTCTGAA ATGATTATT GTTAGCCTTA GTAAAAGATG  
 47001 TATTACCCA TGCTCCATCA TCTAAGGTAT ATTGTAACA AAATGAGAAA  
 47051 AGGTAACCTC ATTTAATGA GAAGAAAAGC AAAATACCTA CATTAAAGTAC  
 47101 TTGAGTCTAT TTAATGTCTG TTAGGGCAGG AAAAAATGGT TATTGCTTT  
 47151 CATATTTAAA ATATCAGCTA CACTCTGGT ATAATATTAA TGGTTGCCAT  
 47201 TTTGACCAGT TTTGTTTAGT GAATAAAAAT TATGTGATTA TTGATCTTAA

FIGURE 3, page 15 of 27

47251 AAAATGTAAT ATCAATTAAA AGGAAAGGAC AGACTCATT TCACCAAAGT  
 47301 AGCAAGTATT TATTAATGTT CCACTTCTT TTTAGCATGG TGCTAGATAAC  
 47351 AGTCATAAT ACAAAAAGAA CATGGACCCA ATCTCGACTC TAATCAAGTT  
 47401 GAGGAGACAA GATGAACACT GAGAATACAA TAGTGAGGAA TACTAACAAA  
 47451 TATATACAAG GTTAAAAGAG TCTAAGTATG CTAGGAATAT AGGGGAAGAA  
 47501 AGAGCTGAAG TACTTCAGGA AGAGTAGAAC ATGAGGCTTT ATTAAAAGA  
 47551 TTAGCAGAAT TTAAGGAAAA GGTGACTTTG TTGAAGATTA TAATGTGAAG  
 47601 ACAAAAGAAC GAGGATGGGA ATAAATTGTT TATTGATGAG GCTTGAAAGA  
 47651 AATTGACTCT AGAGAGTATA TTTGGGTAC TTTGGGAAA TGAAGTTGGA  
 47701 TTAGTGAGAA CGAACAGATT ATGAAAAGAC AAGAACCTG ATTAATGTCA  
 47751 GGATGATTTT ATATTTGAAG TTGGTCAGAT TTATGGCAGT CCTGGCTTTG  
 47801 CCATTTTAG TTTGATGACT TTGAGAAAGT TCCTTCTTGA AGTTTTAATT  
 47851 TTCTGTATAT AAAAGTAAT AACACCTGGT GATCTGCTAG GTTTGTGTTG  
 47901 AGGATTATAT GAGATAAAA GCATGCAAA CTGTTATAAT AGTGCCTGGT  
 47951 AAAATAAGTG CCTAGTTTA AAAACAAGTC TTTGTAAACT GCTTAGGACA  
 48001 TGCCTGGTAT AGGGTAGGTA TGTAATACAT AGTAGGTAGG ATCTGTCCTC  
 48051 TTGCTATTT TAGGAAAAA AACAAAAGGA AGAGCTTCAG CTTAACACAG  
 48101 TATGAACCTGA CGAGCCCTGG TAGGTTTTG ACCAAAAGAG CAACACAGTA  
 48151 AAAGTAGTAC TTAGGAAAGA TTAACAAGGG AACATGGCTT ATACAGTGGT  
 48201 AATGGGGCCT GGAGTCAGG AGGTAAGATA AAATGGTATT ATAATTAAGG  
 48251 AATAGCCAGG CACGATGGCA CATGCATGTA ATGCCAGCTA CTGGAGAGGC  
 48301 TGAGGTGGGA GGATCATGGG AGTCCAGGAG TTTGAGACCA GCCTGGGCAA  
 48351 CTGAGTGAGA CCCAAATCC TAAAAAAATAC AAAGTAAAAAA AGGAATAAAG  
 48401 TCATGAGGGC TTGGACTGGA TTGATAACAG TGAGAATACC GAGAAAGGGA  
 48451 CCATAGGCAG TGTGAACGCA GCTCACTGCA GCCTCAAACC CCAGCCCCAA  
 48501 CGAGCCTCCC ACCTCAGCCT CCCAAGTAGC TGGGACCACCA GACATACACC  
 48551 ACCATGCATG ACTACTTTT TTAGTTTTA CTTTGTAGA GACAGGGTCT  
 48601 CACTGTATTG CCCAGGCTGG TCTCAAACTC CTTGACTTAA GTGATCTTCC  
 48651 TGCCTTGGCC TCCCAAAGTG ATTACAGGCA TGAGGCCACAG TGCTGGCCC  
 48701 AAATAGTTT CTGTGAGTGA ATATTACTTG CATCGTTAAT GTAAATCAA  
 48751 GGCATCAAAG TATTTTACTC TTTTGAAAAA AAATTTAGAG GAGAAATTAA  
 48801 TTATATTAAT ATTCTACCCA TATATGAGTT TAATTTGTAA ATTGTAGCAA  
 48851 AGCATGATGT GCTTACTAA ATTCTTTAT AATTAGAATA AGCTTTTATA  
 48901 AGGGTGAAT TATGTCTTG CTACAGCACT AAACCAAAAT GGCAAAATTG  
 48951 TTTTAGTCGG TAAGCTTTGC TTTTTTAAAAA TATGAAATAA ACAGGTTTTT  
 49001 AAAATGTTAT TTTAATAGTC TTCTCTGTTA TAAACAAAGA AAATTGGTGT  
 49051 TTCTCTAGAG CTTATTTAAA GTAGTGATTA TTGCTCTAAA AGAGGAGTAG  
 49101 CAGTTTTAGA TGCTAATGCT TTCCCTGAC TGAGTTCTAT TTGCCATTAA  
 49151 GTTTTAACCTG CCTAGTGCAA AAATTCTAAT AAAATGTAAT GATGAGGATC  
 49201 CTGTCCTTCC TGACCAAGTGG GTGCTTACTT TTTTCAGGTG GTGTGACAAG  
 49251 AATGACTGTC TCCCTGGTGG TTATTGTTTG TGAGCTTACT GGAGGCTTGG  
 49301 AATATATTGT TCCCTTATG GCTGCAGTCA TGACCAAGTAA ATGGGTTGGA  
 49351 GATGCCCTTG GCAGGGAAAGG CATTATGAA GCACACATCC GATTAAATGG  
 49401 ATACCCCTTC TTGGATGCAA AAGAAGAATT CACTCATACC ACCCTGGCTG  
 49451 CTGACGTTAT GAGACCTCTA AGGAATGATC CTCCCTTAGC TGTCTGACA  
 49501 CAGGACAATA TGACAGTGGG TGATATAGAA AACATGATTA ATGAAACCA  
 49551 CTACAATGGA TTTCCTGTCA TAATGTCAAA AGAATCTCAG AGATTAGTGG  
 49601 GATTTGCCCT CAGAAGAGAC CTGACAATTG CAATAGGTAC CCTTTCAAAA  
 49651 ATATATATAT GTATATATGA GATGGATTTC TGGAAGAAAG GAAAGCAATA  
 49701 AGCAGTAACA TTTAATGGGT CGGATTGTTG GGGGCAAGGG ACATTATTT  
 49751 ATGTCCTTAA ACATCTCTG TTCTTTAAGA AAGGAAGGTA TGCTTCAGTG  
 49801 GATGATTTC TGCTATATAT CACAAAATCT GTATTTCAAGG TTTGTCTTTT  
 49851 GATCCGGCAT GTACCAGAAA TTGGAGTCAG ATTATTTCC CACTCAGATA  
 49901 AGCCTAGATA AGTTGATCTT GGTTATCAA AACAGCATGT AATATAAGAC  
 49951 CTTAGCTAAA TGCAATTCTAGT CAAATACATT CTTGTATTTA ATAAAGTTGG  
 50001 CTTATTGGAA TACAAGTTAT TGAAAATCTC ATCTTCATCA GTCTCTTCA  
 50051 TATTAGAATA AACTGTTTT GCTTTATCAG TCTTTGGGGT TAGAATTATA  
 50101 ATATTAATTT ATAATATCTG ATTTAAAGTG ACAATCACTG AGATTTTTAT  
 50151 TTCTGATCAA ATGCCAGGTT GAAAAAGTAT AACGTATCAG TCCTGTTGTG  
 50201 TTTTATGCAAG ACTTTCTGAA AAATACTGTT TAAAGTATT AGCCATAGTG  
 50251 TATTCTTGG AGATAAAATTA AACTTTCTAT AGTTCTGTTT CTCTAAAATT  
 50301 TGTTTTCTC TTTACCTTAT AGTCCCAG TATTGATGAG GAGACCATTAA  
 50351 AGACTTAATA TTTTTTGAC ACAATCTTAT ATCTCTTCA CCAACCCCTA

50401 AAAAGTGACT GAGGATAGGT ACATCAAGCC ATTGCTTGT TACTCCCCAG  
50451 GTTTTAGTGC CAGACCCCTGA ATGGAAGTGT CAAGCCTTTG GCCTGTCCTGA  
50501 AAGGTCAATTC CTGTGAGCAT ATCATCTCCC TTCCAGCTTA CCTCTGTGGC  
50551 CATTGCAAAA GGATTTAAAA ATAATTTTTG TGCCATTGAA ATGGCACAAG  
50601 ACCAGACAGT GTATGTGGGG GAGTGTTCCT CAAATCAAAC TGGAAACTCT  
50651 TTAATTGTA AGAACATTAA AGCAGAGAGA GAAAAAAAGAA AGGAAAAGAA  
50701 AAAAGATCCT ACAGAGAACAA CCCTGTTAG TTTGGGAACA GGCTACAGCT  
50751 TTGGATTTT CAAGGCCCTAG CATTCCCCTC ATTCTAAATT TTACTTAGCT  
50801 AATACAATAG TAGTTGCCAG AGCTGATGAC ATAGTATTTT GTCATGCTTG  
50851 GCTCCGTTCA AGCATTTCAG TTTTTAGCC ATTACCATGG CTAGACCCAG  
50901 TCAAAAAGAAT TTTCATTGTT TAAGATTCCC ATTATCCTAG TTTTTACTAG  
50951 TAGCCAGCCA AAGAAAAGAA AAAGGAGGTC AGAATTTCGG TATTTACATA  
51001 GAAATTTAAG GGGAAAAGGC CAGGCATGTT TTTAAAGTGT GGAAATTAAG  
51051 AACTATTCA TATCCCCTAG ATTGTGTGGA TGTGTTTTT AAAGTTTGT  
51101 TACTGTCTTG AGAGAGAGAA TATTGAGATA GGACATAATG TTGGTTTAAG  
51151 GGAATGAGGG TACTTTCTGT AGGTGAGGTG CCAAGCCATG TCATCAGAAA  
51201 TGTTAGTCAC ATGACTTTCT AAGCACACCT TAAATGTTT ACCGTGTATG  
51251 TTTTTGTAAT GTTTAAATT TTTAACTGGG AAAAAACAGAC CTGTATATTA  
51301 AGTTTTATAT ATATATATAA ATTTAAAATT ACATATATAT GTTTATATAT  
51351 GTAACTTTTA TATGGGAGAG ATATATATTT CTATATCCTC TATAAAAAAAA  
51401 CATATCTATA TATGAAAATT ATGTACGAA ATGTTAATT ATAATTAATT  
51451 ATATAAATAT TAACATAATT ACATTATATA TATAGAAAAC CTAGTGTACA  
51501 GATCTGTATA TAAATTAAAAA ATGTATGTGT TATATATAGT TACATCATAT  
51551 AATACATATA ATTGATATAT ATAATGATAA ATACTTTATT GAAGGATGAA  
51601 AAAATTTCCA TGCTGTCTCA TAAAATAAGA TGTTGACAT ATGCTAACT  
51651 AGATAGATTC TCCTGTTCA TACTAAAGCA GAATGTTGTA AAATATTA  
51701 TCCAAATGAG ATGTCTCAGA TTAAGGCCAT TTCAACAGGA ATGCTGAGAC  
51751 TTTAAAAAAA AAAAAAGTCT GAGGCTGGC GTGGTGGCTC ATGCCTGTAA  
51801 TCCCAGCACT TTGGGAAGCT GAAGCAGGTG GATCACTTGA GGCCAGGAGT  
51851 TTGAGACCAG CCTGGCCAAT GTGGTGAAT CCCGCCTCTA CTAAAATACA  
51901 AAAAAAATAC ATGGGTGTGG TGACGCATGC CTATAATTCC AGCTACTTGG  
51951 GAGGCTGAGG CAGGAAATC ACTTGAACCT GGGAGGTGGA GATTGCAGTA  
52001 AGCCCCACCA CTGCACTCCA GCCTGGCGA AGAGCAAAC CCTGCTCAA  
52051 AAAAAAAAAGCCTGAATT ATATCAGCAA ATGAAAATCTG TAATGTTGTT  
52101 CTCTGTTCA GAGGCCCTTG AATGAATAGC ACTAAAAATA TTTTTTAAAAA  
52151 AATGAAGAAA ATGAAAATTG TAATGTTCT TATTTAAAAG GCCCTTGAAT  
52201 GAGTAGCATC AAAATATT TTAATGGG GGCCAGGGTG GGAGGTTTGT  
52251 TTGGCACCAG GAGATCAAGA CCAGCTTGGG TAACATAGCA AGACCTTTGT  
52301 CTCTACCAAA AAAAAAAAAT TGGGTGTGGT GGTGCCACCT GTATTCTAG  
52351 CTAUTGGAA CACTGATGCA GGAGGATCCC TGGGACTCTA GAGTCCAGAG  
52401 TGAGACCTG TCTCTAAAC AAACAAACAA ACAAAAATG TATTTATGTA  
52451 AAAGTAATAC TTGTTTTTA AATTTTATT ATTAAATT GATAAAAATT  
52501 GTATGTATGT TTATGTGATG TATATATTGT GGAATGGTTA AATCAGGCTA  
52551 ATTAACCTCAG ATTTTTGTG TGTGTGGGG AATATCTAA AATCCCTCTC  
52601 CTTAGCAGTT TCCAAATGAA ATGAAAAGAT AAAAGTGATT TATTTTTTG  
52651 AGACAGCATC TCACCCCTGT TCTCAGGCTG GAATGCAGTG GCACGATCTT  
52701 GGCTTACTTG ATCCTCGACT TCCCTGGCAT CCGGTGATCC TCCCACTTCA  
52751 CTCTCCTAAT TAGCTAGGAC TACAGGCATG CGCCACCATG ACTGGCTAAT  
52801 TTTGTATT TTTGTATAGG CAAGGTTTTG CCATGTTGCC CAGGCTGGTT  
52851 TCAAGCTCCT GGGCTCAAAC GATCCACCTG CCTCAGCCTC CTGAAGTGT  
52901 GGGATTACAA GTGTGAGCCA CCACACCTGG CGAAAAGTGT TATTTTTTA  
52951 AATGACAAAT TTAAGTCAAA GAGATTGAAT GTTCACTTCT GGTACTTTGT  
53001 ATATAAGAGA AACATTCCAT TAAATAATT TTTAAACATT TCTAAAATT  
53051 CATATTTGT CATTAAATGT TTAAACAAATC AGTATAATT CATTGATACA  
53101 GTGTTGTTA TTTGTGCGGT GTTTAAGATT GATAATTGGG GTTAGTTTA  
53151 ATTCAAGAATG TTATTCTATT TAATGTCACA CTTCATGTC TTTTATTGTT  
53201 TATATCTATT AATGAATTAT TTTAGCTATA GTTATTACTG TTTTAGAGAT  
53251 GAGGTCTTCT ATGTTGCCA GGGTAGACTT GAACTCCTGG GCTTCAGCAA  
53301 TCCCCCTCTC AACCTCCGGA GCACATGAGA TTAGAGACGT GTGCCACTGT  
53351 ATCTGGCCTG CTGTAGTTAT TTTTAATTCT TTTGTCTTTC AACTTTATA  
53401 CTAGAGTTAG AAATGATTAA CAAACCTAT TGCAAGTTTA GAGCGTTATG  
53451 AATTGACTA TATATTCTT ATAACAACCTT AACTTCAGTT GCTTACAAAA  
53501 ACTACAGAGT TTTACTCCCC CGTCCACATT TTATACTATT GATGTCACAC

53551 TTTACATCTT TTTATTTGT GAATCCATTA ATGATACTTC TGGTAGTTT  
53601 TACACTCCAC TATTCAAGTT TCAGACACCA TTCAAGTTGTT AGATTGTTAT  
53651 GAGCTAAAAG CAACTTAATG GGTATTTTC AAAAATCATT TATGTCATT  
53701 GCTAATGGAC TTCTTTCTA TGCCATGATC ATGCTTTTT TATTTTGAG  
53751 ACGGAGTTTC ACTCTTGTG CCTGGGCTGG AGTGAATGG CGCGGCCTCA  
53801 GCTCACTGCA ACCTCCGCC CCGGGTTCA AGCGATTCTC CTGCCCTCAGC  
53851 TGGGATTACA GGCATGTGCC ACCGTGCCGG CTAATTTGT ATTTTAGTA  
53901 GAGACAGGGT TTCACCATGT TGCCAGGCT GGTCTCGAAC TCCTGACCTC  
53951 AGTTGATCTG CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGACGTGA  
54001 GCCACTGCGC CTGGCCTGAT CATGCTTTA AGGTGGTTGA GTAAGTACTA  
54051 GTTGTGGGG CTTTACTTAG TGCCCTCCTA CTCAAATGTG TTAGAACATA  
54101 GTTAAGAAGG CTGTAGTGT CAAAAGGAGT AAAAAGCAGT GCAGTGTGTT  
54151 CAGTAATATC TGCTTCTCAA TTTAGGACTG ATGCTTATTA TGGCTTAAAT  
54201 GTTTTGTAG TAAAATTTGT ATTCAAAAAA TATATTTTT TTTCTTTTG  
54251 CGACAGAGTC TTGCTTGTG ACCCAGGCTG GAGTGTGGTG GTATGATCAT  
54301 GGCTGACTGC AGCCCTGACC TTCCGGGCTC AAGTGTATCTT TCCACCTCAG  
54351 CCTCCCAATT ACTTGGGACC ACCAGCATGC TTGGCCGATT TTTTTTTT  
54401 TTTTTTTTGT GTAGAAGCAA GGTTCCCTA TGTTGCCAAG GCTGGTCTG  
54451 AACTTTAGGG CTCATGTGAT ACTCCTGCCT CGGCCTCCCA AAGTGTAGG  
54501 ATTACAAGCC TGAGCCACCA TGCCGGCCA AAATATTTT ACTATAACAA  
54551 ATATCATATC TGTATATACT CAGTTTAAT ACTAACTCAA AGTAGAAACA  
54601 TAAAGCTGAA TGACTATTTT ATTTTCAGAT TCTCTCCATT GAGTTTCTT  
54651 CTCCGTCTTG TGTGATCTCT GAACTTTCT CCATCTTGC CACTTCTTGT  
54701 CTAGCATTTC TTTTTTATCA GCAGTTTCAT TCAGATTTT TTTTTAGTTC  
54751 TTTCAACGGT GGAGTGGAA TAGGCAGCAG GACAGAAGAA CTTGAAGCAG  
54801 AGCACACTGG AGAGGAGAAA TTAACAAAGC CTTTATGAAT AAAACAACCC  
54851 CCCAATATCA GTCTGTGTGC ATTATGAGCA TAATTTGACT TTCACTCAT  
54901 CTGTAATGTT CATGACTTTT CTAGAAAATT ATACTTTAAC ATGAGAAAAG  
54951 AAAAGAACCG AGCTAATTCA TAGGGATGGA GGACACAGCA TAGTCAAAGC  
55001 AAGAATGAAA CTCTCTTCTAG TGCCACCTCC AGTGCAGAAAT AAGTAACATT  
55051 CAGCAGAGGC AGGTTTCATT TGATAATGGA TTCCCTATAAT AACTGCGCT  
55101 CAGAATTGTT GCAGGTTTTA AAATCCCGTA TTCCAAACCC ACTTCCTTAG  
55151 CCCCCAAGTT AGAAAACAGC TTCAAGTAAAG AAAATTGTC GATGATATAA  
55201 CTTTACCAAA AAATAATTTT TTTCCATGAA GATGATATAT TATTGTTGAC  
55251 TTCTAATTCA ATCAAATATA AACAAATGCT AAAATGGCTTT TCAGTTGACT  
55301 CCTTCTTGG TAAAGGAGAA GATAGGAAAA AATGAAGGGAA TCAGAAGTCA  
55351 TAGGATACAT TAATTTTTT TATCTCTGAA TAAACAGGTT GCCTACTTAA  
55401 AAATCTATCA GTTTAAAAGT GTTGGTCTCT TCTCTCTCTT TTCAGAAAGT  
55451 GCCAGGAAAA ACAAGAAGG TATCGTTGGC AGTTCTCGGG TGTGTTTGC  
55501 ACAGCACACC CCATCTCTTC CAGCAGAAAG TCCTCGGCCA TTGAAGCTTC  
55551 GAAGCATTCT TGACATGAGC CCTTTACAG TGACAGACCA CACCCCAATG  
55601 GAGATCGTGG TGGATATTTT CCGAAAGCTG GGACTGAGGC AGTGCCTTGT  
55651 AACTCACAAAT GGGTAAGTCT GGTACCCACAG GAATCAGTTC ACTTGCTAGA  
55701 ATATAGGATC CTTTTAGTG GAATCTATAT AGTTATTAGG GGAGCATGTG  
55751 AGTCAGCTCC CAGGTGGAA AGTCTGTCCT ATGGTATAGT CACAAATATA  
55801 GGATCAGTCA ATCAAATTT ACATTTACTA AGGAATAAGA AAGATGTCAT  
55851 CTGCTGCTC TTTGCCAAAC AGTACATTT GTAAATAATA CCTCAAAGTT  
55901 GGAAAAGAGG TGCTGAAAGA TCTCCAGCAT GAAAGCATGT TGAGCTTAGA  
55951 GTGCTTCTTT TCCTAGGGAA GAGTGGACCT AACCTGCATG GAGCACTGCA  
56001 AAAACCTGTT TTATTTTTGT AAATGTTCA TTTTTAGTAT ATAAATTCT  
56051 AGTACAATAA TAAGTTCTA GATATTTGC TATTTACTCT TTCAGCCAAT  
56101 ATTTGATTTA TCATGTAATG AAGGAAAGAA TATATACTTA AATGAAATT  
56151 GTAAATGAGC TAAAAATCTC CTTAACAAA TGCTTTGTG CTTTTGTCT  
56201 ACCTTTCTCT ATACACAAAT CTTTATATT TATATAACTG CTAAGGACAA  
56251 ATAAATAACTC ATGTATTTAA AATGTATACA TTGATAATTT ATTTTCCAC  
56301 CTTTACACA TGAACGTGCCA GTGTTCTCC ATTGACAGGA ATATAGGAAA  
56351 GAAACAGATG TCACGGGGGT TGTGGAGACCC TTAATGCACA GAATTGATTT  
56401 AGCAAATACA CTACTTCGTC ACCACTGCTC TCTTTCCCTG GACCTGGGAT  
56451 CTGTTCTCTC ACACCTCTT CTTAGGACC CTTCATTTCC ACTATATATT  
56501 CTTTCTTGTGTT GAACCTAAGA ATGTTGTTT ATCCGAAGGC AAATACCAAA  
56551 AAACAGAGGG TATTCTTGGA TTATGCATAA ACTGGATGGC TAATCCTGAA  
56601 CAGCGTAAAG CTGGTTGAAA TTCTAAACAG AGAATCATAG CAGTTTTTG  
56651 TTGTTTTTTT TTTTAACAT GTTGTAGAAA ACACATTGGT GACAGAAATAC

Sequence Data

56701 ATGACTCCTG TCCAGAGAAA GGAGAGAAAA AGAACAGAAA GGAAGGAAAT  
56751 TTGTTTATTG AACACCTTC AATTTCTCA TTTAACTTTG CAGGACCTCT  
56801 GCAAAGTAGG TAGTTATTC CCTACTTTAC AGATGTAGTA ATTAAAGCTC  
56851 AGGAAGCTT ATAATTTGC CCAAAGTCAT GTGGTGAACA AGTCATGTT  
56901 CAAGGAATCA GACTGTCTT CCTACTTTAA AACCCAGCCT CTTGCTACTA  
56951 TTTTGCCTG TAAGTGACTG ATAGAAATCC TCTTCTTTG TGATTTCTTA  
57001 AACTACTAAA ACATTTCTT GGCCAATATA TTAGATTGAG TTAAGAATAG  
57051 AAATATGAAA CTAGAGAATT AGATCTATGT TTAGTGTGTT TCACTGCGCT  
57101 AATTAAAATA ACTCTTCTAG AATATGAAGT AAATCATTAA AGAGATAAAG  
57151 CCCTTAAAGG CAGGGAGTTT AGAATTATTA AATTCTAATA ATTTAGATAC  
57201 TGATTGGAGA AGAGATGTAT TCATAAGTTA TTATTGTTAC TATTGCTCTT  
57251 TGTGTAATAT TGTTGATTA AATGATGGCA CCGACTTCAT TAAGTTAAA  
57301 AACTCAGTAC TAGTTAAATG GGGCAACTTT TCATAAAGCT TTGCTAGTCC  
57351 TTGAGCCCTT TTATTTGTTA AATGGCTCAA CTGGAACCTA AGCTGAGTTG  
57401 TTACAAACTA TTATTTGCTT CAAGTTGTT TCTGTTCTG GCATGGCTT  
57451 TTCTTTGTG TACTGACAAA TATAATGTT ATTCTGTTGA GTTATGGTTA  
57501 ACTATGAACA CAGAACTGTG AGGGATTAAT TTTCATATT CAGTTGTTG  
57551 ATTAATTCCC AGGTATTTGG CAGCATAGAT ATTAGAAAGG AAAATATTTA  
57601 AAAGAAAGTG TAAAAATAAC GAAGTGTATA GAGCGAGGGG TGGATAGCTA  
57651 ATTAAAATTT TGTCTGGTCC TGCCCTGTTCA TATGAAAAAA GGGGTTGGAC  
57701 TTTCTCTAA GGGAAATATAT TAAATTGTT TCATCATATT TTCCTTATTT  
57751 CTGCTCTGTCA AGGAAATATA ATTGATACAT ATATGGGGAG AAAAGAGATC  
57801 ATTTAGGGAA GTGGCTCATG GGACTTTTG TTTTGTGTTA AGTGTATTAG  
57851 GAAGTCGGGT GTTTTTTTTC TCACTTAAAT TATTTAAAAC CCAGAAAAGA  
57901 AATGATATCT TCTGGTTTT AAAGGAGACC ATGAAGTTCT GCATAGCTAT  
57951 CATTGATGTG TAGTTCATAC TGCATTTTA GAAGTGGAAA ATAGTTATTT  
58001 GGAGGAAGAT AACAAATCTG GAACCTTAGG TGCAAGGAGA AAAAGAATAG  
58051 ATGAAAGGGA AAGATGTTTG TAAATTATAA AAATTCAT TAGCTATTGG  
58101 TTTCTGCAC TTTATATTTT AACTGCAGAA TTTTCAAAA TCAGTTAAC  
58151 TTGGTGAAT TAGCAGGATG TTAATAGGAG TGACTCAGAA AAAAACATTT  
58201 TGTACTGTC TAAGTTGGA AAGTATGGA TAAATACAA TTGAGGTTTC  
58251 TTTACTATGG AACTCCTCA GACTTATAAT ATGTTGATAT TCTTTGATT  
58301 CCAGATGAGG GGATGGGTAA TAGGATACAT GTTTTCCAG ACTTGTGTTGA  
58351 AAATGCAACT ATTTTGGGT TGCAGGGAAG GATATAGTAG AACTCATGGG  
58401 AACTGGTGT TCTTGGAAAC TGCTTGGAA ATGCTGGGTT ATGCCCTGTT  
58451 AACTCTTACA TCATTAGTT TTAGCCCCAA AGGAAACAGC AAATAATGTT  
58501 TTATATGAGC CACATTTGC GTTGATTTC CTTCCACTCT GTAAAATTAC  
58551 TAAAGCAGCA CTCTGACTTT ATTATGCTCA AATCGCTCTT CTCCATTAAT  
58601 GTGTGTTTCT CCATCTTTA GGGTTTTAC TTATATAA CAGAGATTAC  
58651 TGTGAAAAT TCTAAATTG CCACTGGTC GTTATACATT TGTAACCTTC  
58701 CTCACAGTAT ATTTTGTGAT TTGGCAGAGT TTACCAATAT AGATGATACT  
58751 AACTGAAATT ATCATTCTG TATAATTGGA TAGAAAAGCA TGAGTAAGAA  
58801 TTCAATTGGT ATTATATTAA ATTAAATTGCC AAGATTTCA CATTTCCTGA  
58851 CTACAACAAT AAAATCAAAT GAATTGATGG CTTAAAAAAA AGAAATCTCA  
58901 AATGTTAGT CAATGAAGAA CATCTATTGA ATGAGTGAAT GTTCATTATA  
58951 TATAGTGCAT TTTCTGAGCT TTTTGGAGG GGGAAAGTTGC TCCCATGCTC  
59001 TGAGAACTTT TAAGGATCGA TACATTATTT TTAACATAAT AATGAGAAAA  
59051 CATGAGCAGA GAACCCATT CTGTCATTCC CATTCTCTAT CCTCCTGCTC  
59101 CCCCACCTCC CACCCCCAGCC ATCAAGCTAA GTAACTATT TACACCTGGA  
59151 CGTAGCTATA GGAACAGGCT ACTTTGAAGT CTCCCTAGTGA CATCCTCAA  
59201 GTCTGAATGT TCAAAGGCGAG TTTAACAGGG AGGTTGACTT AATGAGATCA  
59251 TCAAGGAAAT GTCCAGTCAT CCTGAAGGGT ATTTTGGATG GGCTTCCAGA  
59301 ATTTAAAGAT TAAAGTTTT TTAAGGTTTT TTTATTTCA CTGTTTATAT  
59351 TGCCACATTA ATTTCCATTA TAAAACCACT AACCATAGTT TTGTTTTAAT  
59401 TAGCAATCTA ATTATTTCA TGTATCCTCA TTATGAGAAT TTATGTCCAT  
59451 CACTTGCTT GATGTGATAA CAGTGACATG CTTAAATGAGA AACAAATTGTT  
59501 ATTTAGAAAA AAATGCACAA AGTGAAGTC CTTTAATCC CTAATCATAA  
59551 ATACATTAA TTAGCTTACT TTAAGAAGTG GCAGTCACAG CTCCTGAACA  
59601 TTAGGGAGTG TTTCTTTGG TCAGCATTAT TTATTTAGTG CACATTGCCT  
59651 TTAATTAA TTTGAAATTA TAGTAAATC CACGGGAGTT TTTAAGTCTC  
59701 CTCACAGCCT TTTGCTACCT TTTCACCAAG GTAGATCCAG ATGATAACTG  
59751 CTGTTGTTGTG ACATCATAGA AATTAGAAAA ATATTTCTT CTGAGGAAAG  
59801 AACATTGTAA ATGAAACTCT ACATATCAGA GGTCTATAGC TATGTATCAA

59851 TATTAAGTTT CTTTTGACT TTGCTTGTA GTCATCTTC A TTCCAAACTT  
59901 TCATAATTAT TATTTTTACT TTAAAAAGAA AAATAACCCA CCAATATTGA  
59951 AGATTAGTAT TGTCACCTT TTGAAAGTCA GTAGAATTAA TGCAAAAGGA  
60001 ACCTGGAAC TAAATCATT TTGTTTTAT TTCTAAAGT TCATGAGACT  
60051 CATTCTTATG GTTCATGTT TTATTTTTC TCTCATTCTT TATCATTATG  
60101 ATTGGAAACT CTTTTAATT AATTCTCAC ACAGTTATTA GCATAATAAT  
60151 CTGTTTCAGG ATTGTCTGG GGATCATCAC AAAGAAGAAC ATATTAGAGC  
60201 ATCTCGAGCA ACTAAAGCAG CACGTCGAAC CCTTGGTGAT TAGATATATC  
60251 AGATCTCTC ATTAGACACC TTAGAAGTCA GGAAGCATGA AACTTGTGAA  
60301 CTGTTGAGTT CTGTCCTTCC CAGATATCTG CTGAACAAAA ATATCCTACT  
60351 ATGCTGCCAA TTACATTGT ATCTGATAAA ATGTGTCGT AAGATAAATT  
60401 TAGATATGTG TAAAATCCC TTTATAGAAA GTAAGCAAAA GTTAACATCT  
60451 CTCATCAAAT CATTCAATTAC AATTTCAGAA CTGTAACACAG TTTGGTAGTG  
60501 GAATAAGTGA ATATTATTGG ACATTCTTAA AGTGAATATG GCAAATCTGT  
60551 CTACCTCAGT GGATACACCG GTCTCAGAAG ACACCTGACT GGTTAAAAT  
60601 GTCTGACCCA TCCCCGCAAG CCCTTTTTT TTTTTTAAA TGTTCCCGA  
60651 TCTTGTGGTA GTCTTATGGT AAATCTAACG TCCTAAAGGA TTTAAAGGA  
60701 GCTTAGCAAT TAGAACTGCT TACAGTTAAA TGAGTTTTT AATGGGCACA  
60751 CTAACTAGAG TGTAATGTGT ATATTATTG TGATCATAGC ATTGTTCTT  
60801 TTTCTGCTAT ACCCTGCATA TCTTCAAAGT CACAGTGTGT GTCCTGCCAT  
60851 CTCATTAGTG AATTGTACCT AGATTATGTG TGTGCCCTT TTGTATGATG  
60901 TTTCTGGAAC GCTATAAGCA GCTTTAGAG TCAAATGCAT TCATTTAAC  
60951 TGGCTTTATG TCCTAGTGGT TTCATGACTA CAAATTGAA TTATCTTACT  
61001 GCATAACATA AAAATGTCT GGCTTAGCA ATTAATGCC GAAATTATTT  
61051 TGCCCTGCAA TTGTCATACC TGTATGAAAC CTGTCCTCAGT TTGCTTAAGT  
61101 GCACAACGTAA TTATGTATTCT CTGTTGTAT GCTAATATTT CACAAGTGT  
61151 TCATGCATCC TTTTTAAAAA AACTACTAAC CAGAATATTA TCGTAGCTAC  
61201 TCATTCATTC TGCTTTCTGC TTCACCTATA ATAATCTTT AGGACTGCC  
61251 TCTGATTTT CACCTATCTT TTAATGTAAG CATTAACAAAC TAAGACTTTC  
61301 ATAAAAGCAC TGTATCTAA CTTTCCTGGC CTAATCAAA AAAAGGAAAA  
61351 CATTGATAAG TGCCTAGAA ACTTGGATTCT TTGTTATAGAT TTGTTCTGG  
61401 GGCTCTGATG TTTGGGATTG ACGTTCTGTG CTGACCATT TATATGCATT  
61451 TTATCTTAAT AGTATGTGCT TTCATGAAGA TTCTGATACA AGTGGGCAAT  
61501 CCTTAAATTAA TCTTGGAAA ATTGGTTAAT TTGGTTAAA AAAGGGAAAG  
61551 TGGCTGGGTG CAGTGGCTCA CGCCTGTAAT CCCCAGCACT TTGGGAGGCC  
61601 GGGACGGGTG GATCACAAGG TCAGGAGTTG AAGCCCATT TGGCCAACAT  
61651 GGTGAAACCC TGTCTCTACT GAAAATAATT GGGGCATGGT GGCACATGCC  
61701 TGTAATCCCA GCTACTTGGG AAGCTGAGGC AGGAGAATTG CTTGAACCGG  
61751 GGACCCAGGA GGCAGGAGTT GCAGTGAGCT GAGATCGCGC CACTGACTC  
61801 CAGCTGGGC TACAGAGCGA GACTCTGCT CAAAAAATAA ATAAATAAT  
61851 AAATGAAAAA GAGAAAATAT TGAGAGGATT TGTCATCAT TTTACTGCTC  
61901 TCTTCATGTG ATGGAAATCA ATTTCCCTTC TCAAATGGGA TCAGTATCAT  
61951 TTCCTAGTCA TACATCCATC CAGTTTTGT TACTTTTTG TTGGCATAACA  
62001 TTAATCAAAA TAGCTCTGCT TCATTGAGGC ATGCAGTCCT CAGACTCTCG  
62051 GTGGAAAGGC TGTCTACTA TTAGTGACCA TAGTAACCTT TTATACCAAA  
62101 GGATGGTTGC TGGATAATTAA TAATATCTTT ACCAATAAAG TACTTTTGG  
62151 AAATACAAAAA TCAGGCTGCT TGCTTGCTC TATTCTGTCA AACAAAAGG  
62201 ATTTAGCTAT AGATTTAGCT TCTCCTTTA TTTTCCCTTT TATTCATAG  
62251 GAGTCTTCTG TTTATCCCT TCAGGCGCCT CCTTGGCATT ATAACAAAAA  
62301 AAGATATCCT CGGGCATATG GCCCAGACGG CAAACCAAGA CCCCCTCTCA  
62351 ATAATGTTCA ACTGAATCTC ACAGATGAGG AGAGAGAAGA AACGGAAGAG  
62401 GAAGTTTATT TGTTGAATAG CACAACCTT TAACCTGAGG GAGTCATCTA  
62451 CTTTTTTTTC CTCCCTTACA AAAAAAGAAA GGAAATATAA AAGCCGGGTT  
62501 TTTGCAACAT GTTTGCAAA TAATGCTGGT GGAATGGAGG AGTTGTTGG  
62551 GGAGGGAAAG GAGAGAGAAG GAAAGGAGTG AGGTATTCTC CGTCTAACAG  
62601 AAAGCAGCGT ATCAACTCCT ATTGTTCTGC ACTGGATGCA TTCACTGAG  
62651 GATGTGCCTG ATAGTGCAGG CTTGCCCTC AACAGAGATG ACAGCAGAGT  
62701 CCTCGAGCAC CTGGCTGTG GCTCCAACAT TGCAAAGACA CATTATCAGT  
62751 CCCTATTCTC AGAGGGATTA CTTTGAATTG AGCCATCTAT AAAACTGCAA  
62801 GGTCTTGGCC TTTTTTTAA TCAAAACTGT TCTGTTAAAT TCATGAATTG  
62851 TATAGTTAAG CATTACCTT CTACATTCCA GAAGAGCCTT TATTCTCTC  
62901 TCTCTCTC TCTCTCTC TCTCTACT GAGCTGTAAC AAAGCCTCTT  
62951 TAAATCGGTG TATCCTTTTG AAGCAGTCCT TTCTCATATT GAGATGTACT

63001 GTGATTTAC TGAGGTTCA TCACAAGAAG GGAGTGTTC TTGTGCCATT  
 63051 AACCATGTA TTTGTACCAT CACTAAATGC TTGGAACAGT ACACATGCAC  
 63101 CACAACAAAG GCTCATCAA CAGGTAAGT CTCGAAGGAA GCGAGAACGA  
 63151 AATCTCTCAT TGTGTGCCGT GTGGCTCAA ACCGAAAACA ATGAAGCTTG  
 63201 GTTTTAAAGG ATAAAGTTT CTTTTTGTT TTCCTCTAG ACTTTATGGA  
 63251 TAATGTGACC GGGTCTTATG CAAATTCTCT ATTCTAAAA CTACTACTAT  
 63301 GATATACAAG TGCTGTTGAG CATAATTAAA TAAAATGCTG CTGCTTGAC  
 63351 AGTAAAGAGA AGGAAGTATT CTGATTAGCT GTATCTGGTA TTAATTGCAT  
 63401 GTTAAAACAC TGGAATTTT AAAATTGAAA TTAGATCAGT CATTCTTTC  
 63451 TTTCTCAAG ATATCTCATG GCTGACACTG AAGAAGAAA GTAATTCTATA  
 63501 ACTTGCACTA AATGTATATT TTTTTCTTA AAAATTTAC ATTCTTATT  
 63551 ATATTTTAT GGATTAATAT TTATAAAA CAGATCAGTT AATATTGCAC  
 63601 TTAAGTAATT TTACCTTTT AATGTGATT TTATAGAATA ATTCAAGACTT  
 63651 ACAAAATACAG AGATATGAAC AAAGTTACA GTGGGAACAA AGGTTTAAAA  
 63701 AAAGGTTGTG GTTCTCTCTC TGTGATCCAG TGTGACATA AACCTTTCTC  
 63751 TGATCTTCA CTGCCATCCT CTGGATTATG TCTTCTGACC TGTCCATT  
 63801 GACCCATTAA CTGGAAAGTT GAAAAACTAC ATTAACGTGGA AAGTTGAAAA  
 63851 ACTACATTAC TTTGGAGAA AAAACCGAAA GTTCTGTAT ACCTCTTAA  
 63901 AAAAAAAATC AAACCAAAAAA TGTGAAAACA ATAGAATTGC AAAGATAGCA  
 63951 GTTAAAATTT TAATCTGAAA ATAACCTTTG AATCTCGGGC TAGGTTATGT  
 64001 CCATATTGAGTGGTCAGT GATGGTTGAG ACATTTTTG CAGGATGAGT  
 64051 TAAAATGCAC TGGATTATAT TTGGGATT TTGGGGATTGGA ATTGTCTGTT  
 64101 TTAATCACAG CCTTAATTCA CAATTGGCAA AGGCAGTTA CTCAAAGGAC  
 64151 TGGGCTAAAT ATTCTGTAA TATGCATTT TGATAGGAAA ATGAAATT  
 64201 TGAAACAGA CATTCTCTT TTTTTGGCT GGAGTGCAGT GGGGCATGGT  
 64251 CTTGGCTCAC TGCAGCGTTG ACCACCTGGG CTCAGTGAT ACTCCCACCT  
 64301 CAGCCACCCA AGTAGCTGGC ACTACGGCA CACGCCACCA TGCCCAGCTA  
 64351 ATTTTTTGT ATTTTTAGTA GAGATGGGT TTTGCCATGC TGCCCAGGCT  
 64401 GGTCTCAACT CCTCAGCTCA AGCAACTCTGC CTGCGTGAGC CTCCCAAAGT  
 64451 GGTGGAATT CAGGCCTGGG CCACTGCGCC TGGCCCCAGAC AGACATT  
 64501 TGAAACACAA CTGGCAATGA GCTGTTTTA CATTGAAA GTGATTCTC  
 64551 ACTTCCTAGT CCTTAATTAT AGTATAACCTA TTAAGATCTG TAAGATCCTG  
 64601 AAGACATAAG ATCATGAAGC CATATAAGAA TGAGGATTGA AAGTTGAGCA  
 64651 AAATTTCCGG GATTTGGGA AACATTCTTA GCTGTGCTAT CTGCCTAAAA  
 64701 TTATTCCTTA TTACTCTCTC CCTTTGACAG ACTTCAAGTT TTCTTCATAG  
 64751 CCCTTCAAA GTTTTTGAG CCATCCAGAG TAAATCATT TCTAAATGAT  
 64801 AGTTCTGTAT ATCTCCAAGT CGTCTTAAGT GTATTTGCCT GTGTGCAACG  
 64851 TATTGCTAGA CTATGAACCT CTCAGCATGG CTGCTGGATA ACTTAATTG  
 64901 CCTGAGTTAA TAGCCTTCAA AGGACAAATC GTTCTTTTG CAGATAGCTT  
 64951 CGTAAAACCTT CACATGGAGT TTATTTTATC ATATTCCTT TTTTTATTTC  
 65001 TGCTCCTCCT TTAATTGCC ATCTTGCCTC AGAGACTGAC ATTTCAGGGT  
 65051 GGATATTAAT TAAAGCTTA ATTTGTTT TTGGTATATT TCTATCCCTA  
 65101 GTATTCTAT CTTACTGCTA AAATACAGGA AAAAGTGCCTG ATTTTTAATG  
 65151 CATTAGTGG TTTTCTTTGG TGTTATCTGT TCCATTTTC TTTTTCTAC  
 65201 ATTGAAGTGT GTCTCCTTT CAACCAAAAT AATGAAATAG TGGAGACCAT  
 65251 GAAATTGTTG TGCCTGGCTA ATTGGCAAAT TAATTTACCA ATATAATAAG  
 65301 TGTAGCGCCT TGTGTTGAATA CCCTTTGTA GAAGGTATGA TGAGAATGGG  
 65351 CAAGGGTGT (SEQ ID NO:3)

**FEATURES :**

Start: 2159  
 Exon: 2159-2237  
 Intron: 2238-22041  
 Exon: 22042-22199  
 Intron: 22200-30359  
 Exon: 30360-30459  
 Intron: 30460-31475  
 Exon: 31476-31663  
 Intron: 31664-32964  
 Exon: 32965-33087  
 Intron: 33088-34548  
 Exon: 34549-34755

Intron: 34756-37975  
 Exon: 37976-38056  
 Intron: 38057-39552  
 Exon: 39553-40098  
 Intron: 40099-46366  
 Exon: 46367-46553  
 Intron: 46554-49237  
 Exon: 49238-49636  
 Intron: 49637-55445  
 Exon: 55446-55662  
 Intron: 55663-62274  
 Exon: 62275-62362  
 Stop: 62363

**CHROMOSOME MAP POSITION:**

Chromosome 4

**ALLELIC VARIANTS (SNPs):**

DNA

Position	Major	Minor	Domain
1275	T	C	Beyond ORF(5')
1456	T	C	Beyond ORF(5')
5893	G	A	Intron
6226	A	G	Intron
8866	T	G	Intron
10397	C	T	Intron
10621	T	-	Intron
19651	A	G T	Intron
19891	T	-	Intron
20272	C	A	Intron
20412	T	A	Intron
23340	A	G	Intron
29948	T	A	Intron
33579	A	C	Intron
40762	G	A	Intron
40936	T	C	Intron
45998	A	G	Intron
47771	T	C	Intron
48117	C	T	Intron
54563	T	G	Intron
58735	C	T	Intron
59643	C	A	Intron
61638	G	T	Intron
63291	G	C	Beyond ORF(3')
63463	A	G	Beyond ORF(3')
63636	G	A	Beyond ORF(3')
63998	T	C	Beyond ORF(3')

Context:

DNA

Position

1275 GCATTTCAGGAGGAGAATCTCCAGTCTAGAGGAATCCTCTCAGAGGTAGCTATAAAATA  
 TTGAACCTGATCTTCATAAGCATTGTGCGGTTTTGTTTTGTTTTAATGACAGTTT  
 TAAACAAGAAAAGTTGCTTTATTTCTGAACCTTCATAAAAATTTCTATTAAAGAGACAATTT  
 CTGAATTTATAACAATTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTGGCT  
 AAAAGTTAAAACACAAAACCCTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAG  
 [T, C]  
 CCTCTGCTTAGTAAACCTTTTGCGTAGTTAGACACATACAATAGTAAAGTTACTT  
 AGTACGTTGATAGTTCTTCTCTCAAAGCTACAATGTCTTAAGCTAGTTCCCTC  
 AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTAGTGGGATAAAACACTATTCAAC

	TCTTCAGTTATCGGTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTT TTTTTATTTTAATAGACTTAGAACCAAGTTCTGAGAAACCTTGCGATATTGTAGTT
1456	TGAATTAAACAATTCTAGAACAGTTGAGTACCTCACTTGAGACACATTTGCTA AAAGTTAAAACACAAACCCCTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAGT CCTCTGCTTAGTAAACCTCTTTTGCCTAGTTAGACACATACAATAGTAAAGTACTT AGTACGTTGATAGTTCTTCTCTCAAAGCTACAATGTCTACTAGCTAGTTCCCTC AAGAAAGGAAACAAGAACGCCGTGGAGGAGATTGGTAGTGGATAAAACACTATTCAAC [T, C] CTTCAGTTATCGGTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTT TTTTTATTTTAATAGACTTAGAACCAAGTTCTGAGAAACCTTGCGATATTGTAGTT TTTTATGGCTATGACTCACATGACATTACTGTATAAAACTAGTACATTCTCGTAAAC CACACAAACTACTAGAGTGCTCTCATTTCTACATTAGAAATGAAAAGGGCATT GTCTGCATTCAAATTCTCTTTACATCTCTGTATTACTTTCCCTTATATTATC
5893	TCTAGTTGACAAGACTGAGGTAAAGGAATTGTTAAGGAAAAGTCAGAATTCCATCCAGATA TTTGGCTCATACTTTAATCATGAGGCTAAACTGCTTCTCTACACGTATCTTCATAGTA ACTTGTGTTTAAGTCTGGTAGAACGATAAGAAGTTAAACACAGACAGAACCTGTGGA AGTTAGTAAATTCTAGTGAACGATAGAAATGATAGAAATCTCTTCTCCCCAAAGTCC CAAGAACAGATTAGTCTGCTTGACAAGTGTATCAAAGTAGACTGTTCTCACATACAC [G, A] GGGGACTCAATAGGGCATTCTGGTGGATATAATAAAATGAGTAAATGCGATAACAGGAG GAAATGCCTAGTGTGCTCTGGATTAGTTGATAACAACAAAGGCAGCTTGTG AGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGAAGTGGAGGTGGCAGATCCTAGAGGA CTAATGATGGCTTAACCACAAAAAGTGTGCTTGCATTGAA
6226	ATAAAATGAGTAAATGCGATAACAGGAGGAATGCCAGTGTGTTGCTCTGGATTAGTT TTGATACAACAAAGGCAGCTTGTGAGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGA AGTTGGAAGAGTGGCAGATCCTAGAGGACTAATGATGGCTAAACACAAAAAGTGTG CTTGCCTATTGAAATAAAAGTTGGGTCTTATTTTCAATTCTCCCTGAAATTATT TCITGACATTGACTAGTCAGCAGTGTATCTAAATAAGCTTTTGTTGTTCTATTATA [A, G] TAGAGGTTGTTCTTTCTTCCCTTGAAAAGTATCATTGTCACATTATTGAAA ATCCAGGTGTTATATGATATTCTTATTGCCAGAGGGACATTCTGCAGGCTCTTGAAAA TGATTTTAGGATTCAAGATACTTATATATTATTGCCCCAATATTATCCAACTAG AAAATTAAACCTCTCTAAATCCATCTAAAGTGTGTAATTAAAGGAAACAC TAAAGATTCTTATTGGTGTCAAGAAACTCCTGTTCTACACAGTAGTATAAAACAAA
8866	ACATGTAACCAACAATGAAATTATTTAGTGAATTGAGAACTCAAAGTGTAGAGTTGA ATCCCTGTTCTACTACTTGCTAGGGTGTGACCTGGGCTGTTAACCTTGACACCTT GTTTCCAAATTATAAAAGTGGAGATAATAATATCTGTACATTGTTGTTGAGGAT TATATGAACTAATATATGTAATGCTCTGAGAACAAATGTCGGTACACATTAAAGTAA AAATTAGCTGTTCTACTGTTATTAGACATGAGCTAGATAACAGTGGCCTACATG [T, G] GAAAGATTATTTAATTCTGATGTAGTTCAGTTATCTATTGTTTATTGTCCT TTGCATTGATGTCATATCTAAAAACCTGCCTAACTCAGGATCACAAAAATTACTCCTG TATTTTATAATTAGCTCTTAGATCTAGGATCCATTAGCTAATTGTTATATATGG TGTGAGGTAGGGTACGGTTCTACCTGGTGAATAGCCAGTTGTCCCAGCATCA TTTATTCAAAGACTATTCTTCTCACTAGAAAAAATATTCTTAAAGAATAATGAAT
10397	CCAGGCTCCCTGAACCTGGCTCAGATGATAGCCTCTGCCACAGCGCTCTGATT AGCTGGACTACAGGTGTGACCAACTACACGGCTTCTGATGAAATTAAACCC AAATATTGAGCAGAAATAATAGCTTGTGTTATTGTTTCTACTATCTGTCAAGTATA GTATTAAATGTTTACATAATTGCTCCAGTCACATACAATACTCTAGTAGAAGTGG TAACAAAACCAAGGTACTCAAAGAGGTTAATAAGTAACCTGCCTGGATCACAGAACTAA [C, T] GGGAGGCAGGGCTGGAATTGACTCTAGGTCTTCTGACCTCAAAGTGCAGTAAAGTCAT GGAATTCTCTACTAGGCCACCTGGAAGAAAAGTGTATCTTTCCAGTCTTTGTTA CTGTTTCTAGCCAGGAGATAGTAGAGTTAGGTAGTAAAGTAGTACTGGCATTCCGG TAGTCAGCCCTCCAAAAAGTTGATTGTTTGTCTAAACTGGAG CTACTAACCTTCAGGTCACTTCTTATCATCCAAGAGCTGGATATTAGGTAGCAGAA
10621	CTCTAGTAGAAGTGGTAACAAACCAAGGTACTCAAAGAGGTTAATAAGTAACCTGC GC

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.20.501323; this version posted September 20, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

	TGGATCACAGAACTAACGGGAGGCAGGGCTGGAATTGACTTAGGTCTTCTGACCTCA AAGTGCAGTAAAGTCATGGAATTCTCTACTAGGCCACCTGGAAGAAAAGTGTATCTTTT TCCAGTCTTTTGTTACTGTTCTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAG TAGTCACTGGCATCCGGTAGTCAGCCCTCCAAAAAAGTTTGATTTTTTTTTT [T, -] GTCTTAAACTTGGAAAGCTACTAACCTTCAGGTCTACACTTCTTATCATCCAAGAGCTGGA TATTAGGTAGCAGAAACTATGGAATTATCCTAAGTCCTTGAGCTGTTAAA ATTAAATTGGTCTGATTAACACTGTGCTCAAGATTTACATTCTAGGAGCCACAGTTGA TTGGTCTAACTTGGATCTATGTGTTCTTAGCTGGGAGGAGAAGGTATCTGATTGA TACCTTCACCAGGACTGCAGTGAGGGACAGAAGTTCTTAAATAATTGGGTCT
19651	TTTATTTCTGCTACTATGGCAGAATTGAGTTGTTGCAACTGTGTGGCATCCAAAGCCTA AAATATTACTCTCCTGGCTCTTGCAACCCGTTTAGATTATGAGCCTTGGCATTA TTATGTTTGTGTTCTTCTATAGCACACAGTAAGATGTTCTGCCACATTGTGCATAA TTTATGGTTATTCAAGGATTATGCAAGTGTAGCTGCAAGAAAAACCTAGAAGTGA ACTTGCTAGGTTGAAGAGCA [A, G, T] CTGTGTATGTTAAATTGTTAGCTTCGCCTCCAAAGGGATTATTCCATTTCTACT TAAACTACTAATTTGTGATAGGACTCTTCTCCATAGCTTGCTAAATTATGCATTC ACACACTTCATCTTACTAATCTGATAGAGGGAAATGATATTGTGGATTGATTGATT TCTTTTATGTGTTAGCTGAGCTTATTTCATATTAAAAGCCAATTGTATTCTTTT CTTGAGCTATCTTTAATGT
19891	TTTATGCAAGTGTAGCTGCAAGAAAAACCTAGAAGTGAACCTGCTAGGTTGAAGAGCA TCTGTGTATGTTAAATTGTTAGCTTCGCCTCCAAAGGGATTATTCCATTTCTAC TTAAACTACTAATTTGTGATAGGACTCTTCTCCATAGCTTGCTAAATTATGCATT CACACACTTCATCTTACTAATCTGATAGAGGGAAATGATATTGTGGATTGATTGATT TTCTTTTATGTGTTAGCTGAGCTTATTTCATATTAAAAGCCAATTGTATTCTTTT [T, -] CTTGAGCTATCTTTAATGTCTCCTGATACATTCTGAAAGTGTGATACTCATATAA GATATATGGTAACATGTGCAAAGATTATTGACTCTAATGAGGGAACCGCCGTGATG ACAAGGCTGATTGAGAAGAGGATGTGAGATGAAGTGTATATCATCAGTGAAGAAAGC AAATTCTTACAGGGAAAAACAAAACCAACTCTAAGGGTTATTGTTCTACTGGACAG AATTCAATTGCTATTACAGATAAAATTACTATTTCATATTATCTTTACAATCAT
20272	CAAAGATTATTGACTCTAATGAGGGAACCCGCTGATGACAAGGCTGATTGAGAAGAG GATGTGTGAGATGAAGTGTATATCATCAGTGAAGAAAGCAAATTCTACAGGGCAAAAA CAAACCAACTCTAAGGGTTATTGTTCTACTGGACAGAATTCTACATTGCTATTACCA GATAAAAATTACTATTTCATTTTACAATTCTACATTCTAATTTCAGAGTCT ATTCCCTAATCAGTAGTAAATAGCTTCAAAATTCTCCGAGCGTCAGGTGACTATTG [C, A] AGGCTAATTGTTGACACTCGGGCTGACTTTAAGAGAACATGCCATAATTCTTGGCCTT ACTTCAAGTTGGATAATTCTTAACACATTCTCTAATTGCAATTGATTCAAG TGATATTATTCTTTTAAATTGTTACTATTGATCACTTGGGTGTTCT CGGAGAGGGGGATTGGCAGGGTCATAGGACAATAGTGGAGGGAGGTGAGCAGATAAAC ATGTGAACAAAGGTCTGGTTCTAGGCAGAGGACCCCTGGCCCTCACAGTGT
20412	TTATTGTTCTACTGGACAGAACATTCTACATTGCTTACAGATAAAATTACTATTTC ATTATCTTTACAATCATTCTAATTTCAGAGTCTATTCCCTAATCAGTAGTAA TAGCTTCAAATTCTCCGAGCGTCAGGTGACTATTATGCAAGGCTATTGTTGACACTC GGGCTTGACTTAAGAGAACATGCCATAATTCTTGGCCTTACTTCAAGTTGGATAA TTTTCTAACACATTCTCTAACATTGCAATTGCAAGTGTATATTCTTTTT [T, A] AAATTTTTTACTATTGATCACTCTGGGTGTTCTGGAGAGGGGATTGGCAG GGTCATAGGACAATAGTGGAGGGAGGTGAGCAGATAAAACATGTGAACAAAGGTCTCTGG TTTCCTAGGCAGAGGACCCCTGCGCCCTCCACAGTGTGTTGTCCTGGTACTTGAGA TTAGGGAGTGGTGTGACTCTTAATGAGCATGCTGCCCTCAAGCATCTGTTAACAAAGC ACATCTGACCGCCCTAACCCCTAACCTGAGTTGACATAGCACATGTTCAGAGA
23340	TTTTTTTTGGAGGTGGGGACTGTCGCCATTCTGTCGCCAAACTGGAGTGCAGTG GTGCAATCTGGCTACTGCAACCTCTGCCCTCCAGGTTCAAGCGATTCTGTACTCAGC CTCCTGAGTAGCTGGAATTAGGTGTGTCATGCAAGCTAATTGTTGATT AGTAGAGATGAAGTTGGCCATGTCAGGCTAGTCTCAGACTCCTGGCCTCAAGTGA

	TTGGCTGACCTCAGCCTCCAAAGTAGAAAATCTTCTGAAAAATAAAATTCAAATCTC [A, G]
	AAAGGCCCTATATAATTGGTGTGGAAATTACTTGTCAATGAAAATGACTATTTACA CAAATTATAAGCTTCATATTAAATATATGTGTGAAACCTGAAATTCAAATTTTATTA TATTGTTATGAAAGGTACAGCCTGTGAGATTCATCAGATGGTATTACCTTAGGGCAT ATCTAAAATAAAACAGTACATGAAATCCAGTGCTTAATCCAGTGATTCTAAACTT TTGCTCTCAGATCCCCTTAAACTCTTAAAGATATTGAAGAGCTCCAAGGAGGTTG
29948	GACTCTACCAATGGGATCGGAGCTCTCCAAACCTGCATATTAAAAGGCCTATAAGTTTG GGGGTCCCTTGTCCACATGATTCTGTAAACATTTGTATTGACATGGTATTA TTATACACAGATCCTGCTTTAAAGAACATTATAATCCACTTAACGTCTAGGACAGAG AATGACCAGATAATTCAACCATATTGTCTACAGAAGACATATATAAAAGATGGTATGT GTACCAATTGAGGTCAAATTGATTCAATTAAAACAATCTAGGCCAGATTTATAG [T, A] TTGGGACCCCTTGCACTCAAATCTCAAGGTTCTTATTAAAATGCAGATCTGGCTGGC ACGGTGGCTCACACCTGAATCCCAGCACTTGGGAGCCAAGGCAGGTAGATCATTGA GCTCAGAAGTTCAAGACCAGTCTGGCCAACATAGCGAGGGCCAGTCTCATTGAAAGAAAA AAAATTTTAATAAAAATAAAAGCAGATCTGGTAAAGACATGTAGTCTGGTTACA GGTATTAACAACGTCTGTAATGTAGTGATTTGCTCCAGACTTACCTTTCCATTATT
33579	TTACTGTGAAGGCTGATTTTTCTCTCACCACTAATTAAACACATGACTAGGCCAA TTTCAGACTATTAGTTAAACATCAAGAGCCTGGAAGAAGTATCTTGACCTAATGTT CTTGACGGTTAGTTACTTGTGTTATGACCTGAATTTTTTGAGACTG AGTCTTGCTGTCGCCCAGACTGGAGTGCAGTGGCGAATCTCAGCTCACTGCAACCTC TGCGTCCCAGGCTCAAGCAATTCTGTGTCAGCCTCTGAGGAGTTGCAGGC [A, C] CCTGTACCATGCCCTGCTAATTGGCATTTTGTTTTAGAG ATGGGTTTCACCATGGCCAGGCTGGCTCAAACCTCAAGTGTACCCCG CCTCAGCCTCCAAAGTGCTGGATTACAGGTGTGAGGCCACACGTGGCTATGACCT GATTTGATTCACTTACATTAAATTACCTTTGATTAGATAAGTTAATTCTTG ATTTGGCATTATGTTAGGAAAGTAGTTAATCACAGTGGTCAACAGTACAAACTT
40762	ATCCATCACCTCAAGCATTATCCCTGTGTTACAAACAATCCAATTACACTCTTAATTA TTTTAAGTGTACAATTAAATTATTGAATATAGTTCAAAGACTTCTTCATTGACTAG CACCTAGGCTAAAAAAATTCAAGACACCTGGGCTCCTGGGATCAATCACGCATACTGTGTC TCTTGCTCACTCCC [G, A] CTGTCCTCTCTCTTCTCGCTCCTTTCTCTCTGTGGTTCTAGGGTGG TGGCCTCAGGGATTGGATTCTTATATTAGCTCAGGATTCCAAGAGGGCTTTTT AATGTAGCCAAGAAGTCTGCACCGTGAATTGTTATTCTATTGAGGTAGTCAC AGAGGCCGACCACAT
40936	TACAATTAAATTATTGAATATAGTCACAGACTTCTCATTGACTAGCACCTAGGCT AAAAAAATTCAAGACACCTGGGCTCCTGGGATCAATCACGCATACTGTGCTCTTGCTC ACTCCCGCTGTCCTCTCTCTTCTCGCTCCTTTCTCTCTGTGGTTCT AGGGTGGTGGCCTCAGGAATTGGATTCTTATATTAGCTCAGGATTCCAAGAGGGC TGTGTTATGTAGCCAAGAAGTCTGCACCGTGAATTGTTATTCTATTGAGG [T, C] AGTCACAGAGGCCGACCACATTAGAGGAGGGACATACACTTGCTGGACAAGTGTAG AGAATTGATGATCATGTTAAACCACTTTATTAGTTCTATTGCTGCTGTAAATAA TTACCCACAATTGCTTAAAGCCACACAAATTAAATTCTACAGTCTGCAAATC AAAAGTCTGAAACGGACTCACTGTGCTAAATTAGGTGTTCTAGGGCATTCTGGAGG CTGAGGAGAGTCTGTTTTGCTTCTGGTATTAAAGCTGCCAGCATTCT CTGAGGAGAGTCTGTTTTGCTTCTGGTATTAAAGCTGCCAGCATTCT
45998	TGTATATCAGTCAAAATTGGCAACTCTGATAAGTTGTCACCTAACATTGAC TTAAGATGAATAGCATCTACCATTCCGTATTGAAATATAGGAGGACATAATCAC ATAATTGAGTAAAGACAGTGTCTAAACTGAATCAGTTAAGTTATGAAAATAC TTCATATTGACTTTAAATATATTGTTAATTCAATAGCTTTGGGTTACAAGT GGTTTGGTTACGTGGATGAATTCTATAATGGTAAGTCTAAGATTACTGCAACTGTC [A, G] CCCAAGTAGTATATATTGTATCCAGCATATTGCTCTTTTTCTTTTTTCA TTTCACCATGGACTAATGAAAATTGTTAGGGACTGACATTAGGGCACCCCTGAGCTAC CTTGAGCTAAAGGAAATAACCCCTGAAATTCTGTTGGCTAGAGAATGTGGTTG

PUBLICATIONS RECEIVED

	TTTTGTAAGTGAATTCATGGATTGTTAAGGTACAAGATTTGCTTAGTTAGTTGTA CTAGGATTTGCTATATTAACATGTAAAAGAATCAAAGTGTAGAAATAATGCA
47771	<p>GAAGAGTAGAACATGAGGCTTATTAAAAAGATTAGCAGAATTAAAGGAAAAGGTGACTT TGTGAAGATTATAATGTGAAGACAAAGGAACGAGGATGGAATAAATTGTTGATTCATG AGGCTTGAGAAGAAATTGACTCTAGAGACTATAATTGGTACTTTGGAAATGAAGTTG GATTAGTGAGAAGGAACAGATTATGAAAAGACAAAGAACCTGATTAATGTCAGGATGATT TTATATTTGAAG</p> <p>[T, C]</p> <p>TGGTCAGATTTATGGCAGTCCTGGCTTGCCATTAGTTGATGACTTGAGAAAAGTT CCTTCTGAAGTTTAATTCTGTATATAAAAAGTAATAACACCTGGTATCTGCTAGG TTTGTGAGGATTATATGAGATAAAATGCATGCAAACACTGTTATAATAGTCCTGGTA AAATAAGTGCCTAGTTAAAACAAGTCTTGAACTGCTTAGGACATGCCTGGTATA GGGTAGGTATGT</p>
48117	<p>GACTTTGAGAAAGTTCCCTCTGAAAGTTTAATTCTGTATATAAAAAGTAATAACACC TGGTGTCTGCTAGGTTGAGGATTATATGAGATAAAATGCATGCAAACACTGTTA TAATAGTGCCTGGTAAAATAAGTGCCTAGTTAAAACAAGTCTTGAACTGCTTAG GACATGCCTGGTATAGGGTAGGTATGTAATACATAGTAGGTAGGATCTGTCCTGCTA TTTTAGGTAAAAAAACAAAAGGAAGAGCCTCAGCTTAATACAGTATGAACTGACGAGCC [C, T]</p> <p>TGGTAGGTTTGAGCAAAAGAGCAACACAGTAAAAGTAGTACTTAGGAAAGATTAACAA GGGAACATGGCTTACAGTGGTAAAGGGCTGGAGTCAGGAGTAAGATAAAATGGT ATTATAATTAAGGAATGCCAGGCACATGGCACATGCATGTAATGCCAGCTACTGGAGA GGCTGAGGTGGGAGGATCATGGAGTCCAGGAGTTGAGACCAGCCTGGCAACTGAGTG AGACCCCCAATCCTAAAAATACAAAGTAAAAAGGAATAAAAGTCATGAGGGCTTGGACT</p>
54563	<p>GCTTTGTCACCCAGGCTGGAGTGTGGTGGTATGATCATGGCTGACTGCAGCCCTGACCTT CCGGGCTCAAGTGTCTTCCACCTCAGCCTCCAATTACTTGGGACCACCAGCATGCTT GGCGATTTTTTTTTTTTTTGTAGAAGCAAGGTTCCCTATGTTGCCAAGGC TGGCTTGAACTTAGGGCTATGTGATACTCCTGCCTGGCCTCCAAAGTGTAGGAT TACAAGCCTGAGCCACCATGGCGCCAAAATATTTCACTATAACAAATATCATATCTG [T, G]</p> <p>ATATACTCAGTTTAATACTAACTCAAAGTAGAAACATAAGCTGAATGACTATTATT TTCAGATTCTCTCCATTGAGTTCTCTCCCTGTGATCTCTGAACCTTCTCCA TCTTGCCACTTGTCTAGCATTTTTTATCAGCAGTTCACTCAGATT TTAGTTCTTCAACGGTGGAGTGGAAAGTAGGCAGCAGGACAGAAGAACTGAAAGCAGAGC ACACTGGAGAGGAGAAATTACAAACGCTTATGAATAAAACAACCCCCCAATATCAGTC</p>
58735	<p>TGGGTTATGCCCTGTTAAGTCTTACATCATTAGTTAGGCCAAAAGGAAACAGCAAAT AATGTTTATATGAGCCACATTGCGTTGATTTCTCCACTCTGTAAAATTACTAAA GCAGCACTCTGACTTTATTATGCTCAAATCGCTCTCCATTATGTGTTCTCCAT CTTTAGGGTTTACTTTATAAAACAGAGATTCTGTGAAATTCTAAATTGCCAC TGGGCGTATACATTGTAACCTCCTCACAGTATATTGTTGATTGGCAGAGTTAC [C, T]</p> <p>AATATAGATGATACTAACTGAAATTAAATCATTCTGTATAATTGGATAGAAAAGCATGAGT AAGAATTCAATTGGTATTATATTAAATTGCAAGATTTCACATTCTGACTACA ACAATAAAATCAAATGAATTGATGGCTAAAAAGAAATCTCAAATGTTAGTCATG AAGAACATCTATTGAATGAGTGAATGTTCAATTATATAGTCGATTCTGAGCTTTT GGAGGGGGAGTTGCTCCATGCTCTGAGAACATTAAAGGATCGATACATTATTTAAC</p>
59643	<p>GTTTATATTGCCACATAATTCCATTATAAAACAGTAACCATAGTTGTTAATTA GCAATCTAATTATTTCATGTATCCTCATTATGAGAAATTATGTCATCAGTTGCTTG TGTGATAACAGTGCATGCTAAATGAGAAACAAATTGTTATTAGAAAAAAATGCACAAAG TGAAAGTCCTTTAATCCCTAAATCATAAATACATTATTAGCTTACTTTAAGAAGTGGC AGTCACAGCTCTGAAACATTAGGGAGTGTCTTTGGTCAGCATTATTATAGTC [C, A]</p> <p>ATTGCCTTAATTAAATTGAAATTATAGTAAAATCCACGGGAGTTTAAGTCCTC ACAGCCTTGTACCTTCAACAGGTAGATCCAGATGATACTGCTGTGTTGTGACA TCATAGAAATTAGAAAAATATTCTCTGAGGAAAGAACATTGAAATGAAACTCTACA TATCAGAGGTCTATGCTATGTAATATTAAAGTTCTTGTACTTGCTTGTAGTC ATCTTCATCCAAACTTCATAATTATTATTTACTTTAAAAGAAAAATAACCCACCA</p>

61638 AAAAAAAGGAAAACATTGATAAGTGTCCCTAGAAACTGGATTCTTTATAGATTTGTTCT  
TGGGGCTCTGATGTTGGGATTGACGTTCTGTGCTGACCATTATGCATTTATCTT  
AATAGTATGTGCTTCATGAAGATTCTGATACAAGTGGCAATCCTAAATTATCTTG  
AAAATTGGTTAATTTGGTAAAAAAGGGAAAGTGGCTGGGTGCACTGGCTACGCCGT  
AATCCCCAGCACTTGGGAGGCCGGGACGGTGGATACAAGGTCAAGGAGTTGAAGCCA  
[G, T]  
TCTGGCCAACATGGTGAACCCCTGCTCTACTGAAAATAATTGGGCATGGGCCACATG  
CCTGTAATCCCAGCTACTGGGAAGCTGAGGCAGGAGAATTGCTTAACCGGGGACCCAG  
GAGGCAGGTTGCACTGAGCTGAGATCGCCTGACTGCAGCCTGGCTACAGAGC  
GAGACTCTGCTCAAAAATAATAATAAATGAAAAGAGAAAATATTGAGAGGA  
TTTGGTCATCATTTACTGCTCTCATGTGATGGAAATCAATTTCCTCTCAAATGG

63291 GAGATGTACTGTGATTTACTGAGGTTCATCACAGAACGGAGTGTCTTGCCATT  
AACCATGTAGTTGTACCATCACTAAATGCTGGAACAGTACACATGCACCACAAAG  
GCTCATCAAACAGGTAAAGTCTCGAAGGAACCGAGAACGAAATCTCATGTGCGT  
GTGGCTCAAACCGAAAACAATGAAGCTGGTTAAAGGATAAAGTTCTTTTGT  
TTCCTCTCAGACTTTATGGATAATGTGACCGGGCTTATGCAAATTCTATTCTAAA  
[G, C]  
TACTACTATGATATAACAGTGTGAGCATAATTAAATAAAATGCTGCTGTTGACA  
GTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCAATTGCT  
GGAATTTTAAAATTGAAATTAGATCAGTCATTCTTCTTCTCAAGATATCTCATGG  
CTGACACTGAAGAAGAAATGTAATTCAACTGCACTAAATGTATATTCTTCTTAA  
AAATTTACCATCTTATTATTTATGGATAAATTATAACAGATCAGTTA

63463 TGTGCCGTGGCTCAAACCGAAAACAATGAAGCTGGTTAAAGGATAAAGTTCT  
TTTTGTTTCCCTCTCAGACTTTATGGATAATGTGACCGGGCTTATGCAAATTCTAT  
TTCTAAAACTACTATGATATAACAGTGTGAGCATAATTAAATAAAATGCTGCT  
GCTTGACAGTAAAGAGAAGGAAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCA  
TAAAACACTGGAATTTTAAAATTGAAATTAGATCAGTCATTCTTCTTCTCAAGAT  
[A, G]  
TCTCATGGCTGACACTGAAGAAGAAATGTAATTCAACTTGCACAAATGTATATTCTT  
TTCTTAAATTACATTCTTATTATTTATGGATAAATTGCAATTGATTTTATAGAATT  
ATCAGTTAATTGCACTTAAGTAAATTACCTTTTAATGTGATTTTATAGAATT  
CAGACTTACAAATACAGAGATATGAACAAAGTTACAGTGGAACAAAGGTTAAAAAA  
GGTTGTGGTCTCTGTGATCCAGTGTGACATAACCTTCTGATCTTCACTG

63636 TGCTGCTGCTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAAT  
TGCATGTTAAACACTGGAATTTTAAAATTGAAATTAGATCAGTCATTCTTCTTTC  
TCAAGATATCTCATGGCTGACACTGAAGAAGAAATGTAATTCAACTTGCACAAATGT  
ATATTCTTCTTCTTAAATTACATTCTTATTATTTATGGATAAATTATA  
AAATACAGATCAGTTAATTGCACTTAAGTAAATTACCTTTTAATGTGATTTTATA  
[G, A]  
AATAATTCAAGACTTACAACACAGAGATATGAACAAAGTTACAGTGGAACAAAGGTT  
AAAAAAAGGTTGTGGTCTCTCTGTGATCCAGTGTGACATAACCTTCTGATCT  
TTCACTGCCATCCTCTGGATTATGTCTCTGACCTGTGACCTTGGACCCATTAGGAA  
AGTTGAAAACACTACATTAACGGAAAGTTGAAAAACTACATTACTTGGAGAATAAAC  
GAAAGTTCGTGTATACCTCTTAAACACAAAAATGTGAAAACAATAGAA

63998 AAAAAAGGTTGTGGTCTCTCTGTGATCCAGTGTGACATAACCTTCTGATCTT  
TCACTGCCATCCTCTGGATTATGTCTCTGACCTGTGACCTTGGACCCATTAGGAA  
GTTGAAAACACTACATTAACGGAAAGTTGAAAACACTACATTACTTGGAGAATAAAC  
AAAGTTCGTGTATACCTCTTAAACACAAAAATGTGAAAACAATAGAA  
TGCAAAGATAGCAGTTAAATTGAAATTCTGAAATAACCTTGAATCTGGGCTAGGTTA  
[T, C]  
GTCCATATTGAAGTGGTCACTGATGGTTGAACATTGGTGCAGGATGAGTTAAATGC  
ACTGGATTATATTGGATTGGTTGGATTGTCTGTTTAATCACAGCCTTAATT  
CACAATTGGCAAAGGCAGTTACTCAAAGGACTGGCTAAATATTCTGTAATTATGCATT  
TTGATAGGAAAATGAAATTGGTCAAACAGACATTCTTGGCTGGAGTGCA  
GTGGGGCATGGCTTGGCTCACTGCAGCGTTGACCACCTGGCTCAAGTGATACTCCCG